Summer 2020

New Models for Truck Appointment Problem and Extensions

Mohammad Torkjazi

NEW MODELS FOR TRUCK APPOINTMENT PROBLEM AND EXTENSIONS

by

Mohammad Torkjazi

Bachelor of Science
Islamic Azad University, 2006

Master of Science
Sharif University of Technology, 2011

Submitted in Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy in

Civil Engineering

College of Engineering and Computing

University of South Carolina

2020

Accepted by:

Nathan N. Huynh, Major Professor

Robert L. Mullen, Committee Member

Yuche Chen, Committee Member

Pelin Pekgun, Committee Member

Cheryl L. Addy, Vice Provost and Dean of the Graduate School

DEDICATION

To my parents, Gohar Dizijani and Abolghasem Torkjazi and my sister, Alaleh Torkjazi. I love you.

ACKNOWLEDGEMENTS

First and foremost, I want to thank my supervisor, Dr. Nathan Huynh for his advice, guidance, and encouragement during my Ph.D. studies.

I would like to express my appreciation to my dissertation committee members, Dr. Robert Mullen, Dr. Pelin Pekgun, and Dr. Yuche Chen for their valuable opinions.

Last but not least, I would like to thank my family, friends, and the graduate students in our research group for their support during the completion of my dissertation.

ABSTRACT

The problematic issues surrounding gate congestion at marine container terminals have been well documented. Random truck arrivals at maritime container terminals are one of the primary reasons for gate congestion. Gate congestion negatively affects the terminal's and drayage firms' productivity and the surrounding communities in terms of air pollution and noise. To alleviate gate congestion, more and more terminals in the U.S. are utilizing a truck appointment system (TAS).

The first study proposes a novel approach for designing a Truck Appointment System (TAS) intended to serve both the marine container terminal operator and drayage operators. The aim of the proposed TAS is to minimize the impact to both terminal and drayage operations. In regard to terminal operations, the TAS seeks to distribute the truck arrivals evenly throughout the day to avoid gate and yard congestion. In regard to drayage operations, the TAS explicitly considers the drayage truck tours and seeks to provide appointment times such that trucks do not have to deviate greatly from their original schedule. The proposed TAS is formulated as a mixed integer nonlinear program (MINLP) and the model is solved using the Lingo commercial software. Experimental results indicate that the proposed TAS reduces the drayage operation cost by 11.5% compared to a TAS where its aim is only to minimize gate queuing time by making truck arrivals uniform throughout the day.

The second study proposes a novel approach to modeling the TAS to better capture the multi-player game (i.e., interplay) between the terminal and drayage firms regarding

appointments. A multi-player bi-level programming model is proposed with the terminal functions as the leader at the upper-level and the drayage firms function as followers at the lower-level. The objective of the leader (the terminal) is to minimize the gate waiting cost of trucks by spreading out the truck arrivals, and the objective of the followers (drayage firms) is to minimize their own drayage cost. To make the model tractable, the bi-level model is transformed to a single-level problem by replacing the lower-level problem with its equivalent Karush–Kuhn–Tucker (KKT) conditions. For comparison purposes, a single-player version of the TAS model is also developed. Experimental results indicate that the proposed multi-player model yields a lower gate waiting cost compared to the single-player model and that it yields higher cost savings for the drayage firms as the number of appointments per truck increases. Moreover, the solution of the of multi-player model is less sensitive to objective function coefficients across problem sizes compared to the single-player model.

Lastly, the third study develops a truck appointment system (TAS) considering variability in turn time at the container terminals. The consideration of this operational characteristic is crucial for optimal drayage scheduling. The TAS is formulated as a stochastic model and solved using the Sample Average Approximation (SAA) algorithm. Using turn time distributions obtained from actual data from a U.S. port, a series of experiments is designed to evaluate the effectiveness of the proposed stochastic TAS model compared to the deterministic version where an average turn time is used instead of a distribution. Numerical experiment results demonstrate the benefit of the stochastic TAS model given its lower drayage cost error by 3.9% compared to the deterministic TAS model. This result implies that the schedules produced by the stochastic TAS model are more robust and are able to accommodate a wider range of turn time scenarios. Another key takeaway from the experiment results is that the stochastic TAS model is more beneficial to utilize when the ratio of quotas to requested appointments is lower. Thus, in practice, when this ratio is more likely to be on the lower end, drayage companies would benefit more if the appointment schedule adopts the stochastic approach described in this paper.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Intermodal freight transport is the use of two or more transportation modes (road, water, air, and rail) to transport containers from their origins to destinations. The transfer of containers from one mode to another mode always take place at intermodal terminals that are specifically designed for the intended modes and where specialized equipment are used to facilitate the transfer of containers. As shown in Figure 1.1, there are three major types of container terminals: 1) marine container terminal which connects the road/rail modes to the water mode, 2) airport container terminal which connects the road/rail modes to the air modes, and 3) rail container terminal which connects the road mode to the rail mode. The focus of this dissertation is on marine container terminals.

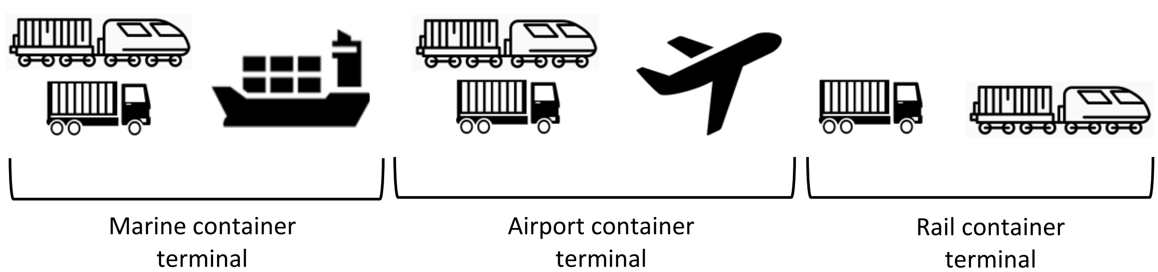


Figure 1.1 Graphical illustration of container terminals

As shown in Figure 1.2, there are three main areas at a marine container terminal: 1) berth, 2) yard, and 3) gate. The berth is a ship's allotted space at a wharf or dock. At the berth, the quay cranes are positioned to load and unload containers to and from the ships. The yard is where containers are stored until they leave the terminal via ship, train or trucks. The containers are either mounted on individual chassis (in wheeled operations) or stacked four-high and six-wide in a yard block (in grounded operations). Wheeled operations facilitate the container pickup and drop-off process, but it requires much more space. Grounded operations allow for more containers to be stored in a smaller area, but it requires the use of Rubber Tired Gantry (RTG) cranes to load/unload containers from/to transporters and road trucks. As shown in Figure 1.2, a transporter is a vehicle that is used to transport containers from the yard to the berth and vice-versa. A road truck is a vehicle that is used to transport a container from the terminal to a hinterland destination (e.g., distribution center, warehouse) and vice-versa. Entering road trucks and exiting road trucks are processed at the gate. Upon entering the terminal, truck drivers are required to present paperwork for their transactions, which could be dropping off an export container or picking up an import container or both. If a driver is dropping off an export container, that container needs to be inspected. Technologies such as high definition cameras and Optical Character Recognition are used to expedite the inspection process. Just like export container, import containers are inspected upon exiting the terminal. In addition, they are screened via Radiation Portal Monitors.

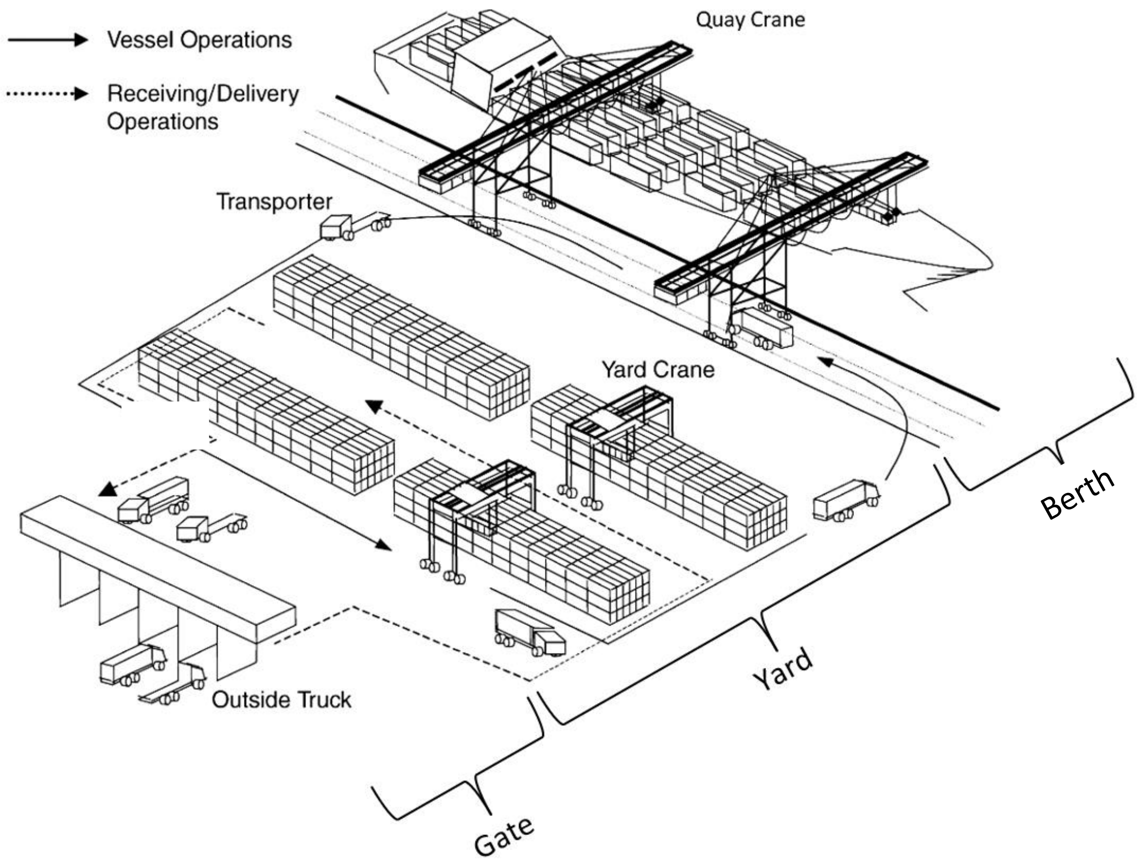


Figure 1.2 Marine container terminal (Park, 2003)

There are several classes of problems at marine container terminals. At the berth, the classical problems include berth allocation, quay crane scheduling, and stowage planning. The berth allocation model provides the location and the duration in which a ship is served (Park and Kim (2003); Heyden and Ottjes, (1985); Lai and Shih, (1992); Brown et al. (1994, 1997); Lim, (1998); Park and Kim, (2002); Kim and Moon, (2003); Guan and Cheung, (2004)). The quay crane scheduling problem deals with assigning multiple quay cranes to load and unload containers into and from the ship; each associated with a start and end time (Park and Kim (2003); Daganzo (1989a; 1989a)). It should be noted that the berth allocation and quay crane scheduling problems are interrelated because the time that it takes to unload and load of a vessel affects the berth time and availability.

The stowage planning problem determines the location of a container group with the same destination to be loaded on the ship (Christiansen et al. (2007)). The objective of this problem is to stack the containers in a manner that requires minimum rehandling in subsequent port visits.

For the container yard, the classical problems include the container stacking problem (for exports and imports), yard crane scheduling problem, rehandling problem, and storage problem. The objective of container stacking problem is to store the containers in a manner such that the total handling cost is minimized (Taleb-Ibrahimi et al., (1993); Castilho and Daganzo, (1993); Kim, (1997); Chen, (1999)). The objective of yard crane scheduling problem is to minimize the waiting time and total distance traveled (Zhang et al. (2002), Cheung et al. (2002), Kim et al. (2003), Lai and Lam (1994), and Lai and Leung (1996)). The rehandling problem deals with shuffling the container stack such that the efficiency of future loading/unloading operations are maximized (Caserta et al., 2011). The storage problem deals with the reservation of space for containers such that the level of reshuffles in the yard is minimized (Jiang et al., 2012).

The gate has attracted much less attention in research compared to the yard and berth. The majority of studies have focused on either quantifying the impact of gate congestion or proposing a method to alleviate gate congestion. The work of Watanabe (2003) have sought to quantify the environmental impact of gate congestion. A number of studies have sought to propose the use of truck appointment systems to reduce the randomness in truck arrivals. These studies mainly deal with the optimization of quotas that determines the number of trucks that can enter the terminal during each time-window (Huynh and Walton (2008), Huynh (2009); Chen et al. (2011); Chen et al. (2013a); Chen et al. (2013b); Zhang et al. (2013)). Other gate-related studies deal with determining the optimal gate layout to minimize total cost (Guan and Liu (2009a, 2009b); Fleming et al. (2013); Minh and Huynh (2017)).

This dissertation aims to improve the design of TAS to improve efficiency of terminals and drayage firms. It develops TAS to consider not only the requirements of the marine container terminal but those from the drayage companies. The first study develops a new mathematical model for the TAS that seeks to minimize the drayage cost when determining the appointment time-window(s) for a truck, as well as the truck's gate queuing time (Torkjazi et al., 2018). To our knowledge, this is the first study to consider truck tours in TAS design. The second study considered the drayage firms and terminal operators as separate entities that have different business interests and needs. Thus, the TAS problem is treated as a multi-player game where each player seeks to optimize his objective function. Lastly, the third study extends the first study to consider uncertainty in turn times.

1.1 LIST OF PAPERS AND STRUCTURE OF DISSERTATION

This dissertation includes three research papers as listed below and they appear as separate chapters:

1. Torkjazi, Mohammad, Nathan Huynh, and Samaneh Shiri. "Truck appointment systems considering impact to drayage truck tours." Transportation Research Part E: Logistics and Transportation Review 116 (2018): 208-228.

2. Torkjazi, Mohammad, Nathan Huynh, and Ali Asadabadi. "Modeling the Truck Appointment System as a Multi-Player Game."

3. Torkjazi, Mohammad and Nathan Huynh. "Truck Appointment Systems with Stochastic Turn Times."

The rest of this dissertation are organized as follows: Chapters 2 to 4 present the three above-mentioned studies and finally, chapter 5 provides a summary of the studies, concluding remarks, and future works.

CHAPTER 2

TRUCK APPOINTMENT SYSTEMS CONSIDERING IMPACT TO DRAYAGE TRUCK TOURS[1]

The problematic issues surrounding gate congestion at maritime container terminals have been well documented. Naboothiri and Erera (2008) reported that gate congestion leads to a decrease in drayage productivity. That is, drivers typically experience longer waiting time when they arrive during peak hours, which would require longer truck turn time (the sum of terminal gate queue time and in-terminal time), and thereby reduce their available time to perform other moves. Truck turn time refers to the time it takes a truck to complete the delivery or pick-up transaction; it is the difference between the gate out time and the gate in time. A byproduct of gate congestion is a concentration of idling trucks. It has been documented that when trucks are idling they emit a greater amount of emissions compared to when they are moving. Emissions from diesel trucks are known to contain a number of carcinogens and are associated with elevated levels of asthma attacks, emergency room visits, hospitalizations, heart attacks, strokes and untimely deaths (Hill, 2005; Sax and Larsen, 2004; Giuliano and O'Brien, 2007; Schulte et al., 2015; Schulte et al., 2017). Heilig et al. (2017a) provided a broad overview of academic works related to

[1] Torkjazi, Mohammad, Nathan Huynh, and Samaneh Shiri. "Truck appointment systems considering impact to drayage truck tours." Transportation Research Part E: Logistics and Transportation Review 116 (2018): 208-228.

environmental sustainability in ports, hinterland operations, and combination of both operations.

To reduce gate congestion, more and more maritime container terminals (e.g., the Port of Baltimore, Port of Vancouver, and Port of Hamburg) are adopting the use of a truck appointment system (TAS) as called in the U.S. and vehicle booking system (VBS) in other parts of the world (Heilig & Voß, 2017). A TAS provides several key benefits to the terminal operators (the company or port authority that is managing the terminal operations). One, it allows the terminal operators to match demands (container transactions) to supplies (labor and equipment availability). Second, it allows the terminal operators to evenly distribute truck arrivals throughout the day, and hence, reduce truck queuing at the gate. Lastly, the advanced entry of container and truck information via the TAS expedites the processing of the trucks upon their arrivals at the terminal.

The typical function of most existing TASs is that they allow the terminal operators to set a quota for the maximum number of trucks allowed to enter a specific yard block or zone during a pre-specified time-window, typically in the range of 1 to 4 hours. Quotas are set based on yard crane availability. They are also set to avoid potential conflicts with other operations in a certain yard block or zone, such as vessel operations, warehouse operations, rail operations, and customs inspections. From the trucker's perspective, once the quota for the desired time-window is reached, he needs to choose a different time-window for the appointment. It is evident that the quotas set by the terminal operators can have a significant impact on terminal and drayage operations. To this end, many studies have sought to determine the optimal quotas for TAS (e.g., Huynh and Walton (2005), Huynh (2009); Chen et al. (2011); Chen et al. (2013a); Chen et al. (2013b); Zhang et al.

(2013)). Other studies have sought to understand the impact of TAS on drayage scheduling (e.g., Namboothiri and Errera (2008); Shiri and Huynh (2016)). However, to date, no studies have developed a TAS that seeks to minimize the impact on drayage scheduling. Given that each drayage firm has a number of timing constraints imposed by customers and the network travel time varies day to day (Torkjazi et al. (2017)), any additional timing constraint imposed by the TAS will make it even more difficult for drayage operators to make deliveries or pickups on time. Thus, an effective TAS must consider not only the capacity and constraints of the terminal but also that of the drayage firms.

The objective of this paper is to develop a new mathematical model for the TAS that seeks to minimize the drayage cost when determining the appointment time-window(s) for a truck, as well as the truck's gate queuing time. The gate queuing time is minimized when the appointment quotas (and hence truck arrivals) are distributed evenly throughout the day. To determine the right balance between reducing drayage cost and gate queuing time, different weights associated with these two cost components are evaluated. To our knowledge, this is the first paper that proposes to design a TAS that explicitly takes into account truck tours. The TAS model is formulated as a mixed integer nonlinear problem (MINLP) and can be solved using the Lingo commercial software. Given the combination of linear and non-linear constraints, and integer and real decision variables, the proposed TAS model is a MINLP.

The rest of the paper is organized as follows. Section 2.1 provides a summary of closely related studies to provide context for the contributions of this work. Section 2.2 provides the problem description and formulation, followed by Section 2.3 which presents the numerical experiments. Section 2.4 discusses the managerial insights. Lastly, Section 2.5 provides a summary of the study and concluding remarks.

2.1 LITERATURE REVIEW

As mentioned previously, a number of studies have sought to develop methodologies to determine the optimal quotas for TAS or assess the effectiveness of TAS. A comprehensive review of TAS can be found in the work of Huynh et al. (2016). The following review focuses on studies that examined the impact of TAS on drayage and truck emissions. Table 2.1 shows summary of TAS literature reviews. While previous studies on TAS have considered truck queuing, emissions, terminal resources, and the impact of shifting a truck's desired appointment, there has been no study that explicitly considered the effect of truck's schedule of jobs in determining the terminal appointment time-windows on drayage operation cost. The notion of considering drayage scheduling in designing the TAS is partially addressed in the work of Phan and Kim (2015; 2016). In the earlier study (2015), the authors proposed an iterative scheme where there is coordination between the TAS and drayage firms. In this scheme, when a drayage firm receives a pickup/delivery order for an inbound/outbound job from a consignee/shipper, the first step is to submit a tentative appointment request for the pickup/delivery of the container at its most desired time-window. In the second step, the TAS estimates the truck queuing time at the terminal gate based on various appointment requests and this information is shared with the drayage firms. Having knowledge of the truck queuing time, Phan and Kim's proposed iterative scheme requires drayage firms to repeat step 1 to submit new tentative appointment requests, considering the truck queuing time, the available number of trucks at various time-windows, and the cost of changing the time-window for the delivery. These two steps are repeated until all drayage firms confirm their appointments. In their latter work (2016), the authors extended their previous work (Shiri and Huynh, 2016) to consider each set of trucks' schedule of jobs. Since a trucking company may have delivery orders other than those to and/or from the container terminal, the dispatcher of the trucking company constructs a truck travel schedule. They proposed a scheduling problem in which a trucking company determines the number of trucks to start a new job after finishing the previous job according to minimize the fixed cost for deploying a set of trucks for the day, the truck travel cost, and truck waiting cost for a trucking company. Their scheduling problem also gives the suitable terminal appointment time for each set of trucks. Similar to their earlier study, an iterative process of submitting the truck appointment application including the number of trucks and the appointment time-window for truck arrival to the TAS by a drayage firm and updating gate queuing time from TAS is performed until all drayage firms confirm their appointments.

The TAS poses a significant challenge to drayage firms and affects drayage scheduling. A few studies have addressed the drayage scheduling problem with time constraints at maritime container terminals imposed by the truck appointment system. The TAS in these studies is limited to a quota per time-window at terminal. Namboothiri and Errera (2008) studied the situation involving a single drayage firm that serves a number of inbound and outbound container move requests to and from the terminal from a single truck depot location. The authors assumed that the drayage firm needs to book an appointment in advance prior to each visit to the terminal. They formulated the problem in two phases. Phase I seeks to maximize revenue by determining which outbound and inbound container move requests should be served given the available appointments at the terminal. Phase II determines the route and schedule for each truck to serve the container moves identified in phase I. The objective of phase II is to minimize the required fleet size. Phase I was formulated as an integer programming model and its simplicity allowed the authors to obtain optimal solutions via CPLEX. For phase II, the authors proposed a heuristic based on column generation that generates near-optimal solutions. Shiri and Huynh (2016) also addressed the drayage scheduling problem with time constraints at maritime container terminals imposed by the truck appointment system, in addition to the time constraints at customer locations. They considered the situation where a drayage firm has multiple truck depots and it has to manage a fleet of trucks to satisfy the container move requests between the customers' locations, the empty container depot and the maritime container terminal. The authors proposed an integrated model as an extension of the multiple traveling salesman problem with time-windows (m-TSPTW) that solve the empty container allocation problem, vehicle routing problem and appointment booking problem in an integrated manner. To solve the proposed model, the authors developed a reactive tabu search (RTS) meta-heuristic. However, the TAS in these studies are assumed to be an online appointment system that put a quota on number of appointments in a time-window. So, none of these studies have solved a separate problem for the TAS to adjust the requested appointments.

Table 2.1 Summary of the TAS literature review

Author(s) (year)	Study country						TAS design								Solution method				
													Queuing system		Simulation model			Opt*	
	US*	CA*	CN*	CL*	FR*	H*	Q*	Y*	E*	A*	D*	C/I*	S*	NS*	AG*	DE*	Qn*	Ex*	Hu*
Morais and Lord (2006)	✓	✓					✓		✓	✓			✓				✓		
Huynh and Walton (2008)	✓						✓						✓			✓		✓	
Huynh (2009)	✓						✓			✓			✓			✓			
Guan and Liu (2009a;2009b)	✓						✓						✓					✓	
Zhao and Goodchild (2010)						✓	✓	✓											✓
Chen and Yang (2010)			✓				✓	✓	✓		✓		✓						✓
Chen et al. (2013a)			✓				✓	✓						✓					✓
Chen et al. (2013b)	✓						✓		✓		✓			✓					✓
Zhang et al (2013)			✓				✓	✓						✓					✓
Zehendner and Feillet (2014)					✓		✓							✓		✓			✓
Schulte et al. (2015)				✓			✓		✓		✓	✓	✓			✓			
Phan and Kim (2015)						✓	✓	✓				✓		✓				✓	
Phan and Kim (2016)						✓	✓	✓			✓	✓		✓				✓	
Schulte et al. (2017)				✓			✓	✓	✓	✓	✓	✓						✓	

Table 2.2 Summary of the TAS literature review (Continued)

Author(s) (year)	Study country						TAS design								Solution method				
													Queuing system		Simulation model		Qn*	Opt*	
	US*	CA*	CN*	CL*	FR*	H*	Q*	Y*	E*	A*	D*	C/I*	S*	NS*	AG*	DE*		Ex*	Hu*
Current study	✓						✓	✓		✓	✓			✓				✓	

*US: United States; *CA: Canada; *CN: China; *CL: Chile; *FR: France; *H: Hypothetical; *Q: Quotas considerations; *Y: Yard considerations; *E: Environmental considerations; *A: Assessment of existing TAS methods; *D: Drayage firms considerations; *C/I: Collaborative/Iterative TAS; *S: Stationary; *NS: Non-Stationary; *AG: Agent-Based; *DE: Discrete-Event; *Qn: Quantitative; *Opt: Optimization; *Ex: Exact; *Hu: Heuristic

The effect of collaboration between trucking companies to reduce emission and increase the benefit of trucking companies by reducing empty truck trips has been addressed by Schulte et al. (2015; 2017). In their earlier work, Schulte et al. (2015) developed a collaborative discrete event and system-based simulation to model drayage scheduling at port of San Antonio, Chile and evaluated scenarios for number of empty trips, emission savings, capacity utilization and costs. They found that a collaborative TAS needs to be integrated in port operations in order to reduce emissions; otherwise, it may worsen port congestion. In their latter work, Schulte et al. (2017) assessed the effectiveness of central scheduling model for truckers compared to a decentralized model of scheduling every truck separately. They developed a graph-based mathematical model based on the m-TSPTW to assign jobs to trucks. Their model optimizes the gain, travel cost, and emissions for those jobs that can be performed by multiple trucks. They found that a collaborative TAS significantly reduces emissions, travel cost and number of required trucks. However, in both studies, they assumed that trucking companies will collaborate among themselves and operate under one TAS. This approach allowed the authors to solve the TAS problem from a centralized perspective considering different trucking companies. Heilig et al. (2017a) proposed an economic and environmental analysis for the inter-terminal drayage problem. They extended their earlier work (Heilig et al., 2017b) by proposing a multi-objective model with the goal of minimizing the variable and fixed travel cost and emissions. They also provided a cloud-based decision support system to support real-time communication between truckers and dispatchers. They concluded that such a system would benefit ports operationally and at the same time address environmental factors.

In this study, we propose a novel approach to designing the TAS. Our objective is to design the TAS such that it poses the least impact to a truck's schedule. Such an approach would benefit drayage firms, and hence, would improve compliance and effectiveness of TAS. We integrate the developed TAS model in this study with the drayage scheduling model from the work by Shiri and Huynh (2016) to examine the effect of the proposed TAS model on drayage scheduling. Our design of TAS assumes that the cut off time for making an appointment is one day prior to the appointment, as often is the case in practice in U.S. (e.g., Port of New York/New Jersey, Ports of Los Angeles and Long Beach). Furthermore, our proposed TAS only allows the drayage firms to submit appointment requests once, and that they have to accept whichever time-window is given by the TAS. For pickup cases, if the container is not available at terminal yet, a notification will be sent to the customer whenever it is available and customer can request an appointment for the next day. Note that the proposed TAS does not allow for negotiation of appointments between the drayage firms and the terminal operator. This approach seeks to provide an optimal and equitable assignment of appointments to all drayage firms. Allowing drayage firms to negotiate appointments is not only unpractical, but also put some firms at a disadvantage. The drayage operation cost is a function of terminal appointment time. So, from knowing all the appointment requests for a particular day and knowing the quotas for each time-window, our TAS seeks to keep each truck's requested appointment time(s) whenever possible. If it is necessary to change an appointment time for a truck that has more than one appointment, our objective is to keep the time-gap between two consecutive appointments for a given truck greater than or equal to the time-gap between the desired appointment requests; note that the TAS knows the time-gaps between a truck's appointments since the truck's ID is submitted with each appointment request. This approach would prevent the situation where two given appointments are too close together, and thereby not provide the trucker with sufficient time to perform the necessary moves.

2.2 PROBLEM DESCRIPTION AND FORMULATION

This study considers a maritime container terminal and multiple drayage firms in a port setting similar to that of Ports of Los Angeles and Long Beach in the U.S. Each drayage firm owns a number of trucks that can be used to perform container move requests, and each request must be done within a pre-defined customer's time-window and appointment time-window at the maritime container terminal.

Figure 2.1 shows the TAS framework utilized in this study. Based on the Port of Vancouver TAS model, it is assumed that on a daily basis drayage firms will have submitted their appointment requests by 5:00 PM for the next day's appointments (step 1). Similarly, it is assumed that the terminal operator will have submitted the quotas for each time-window by 5 PM (step 2). For each appointment request, the information to be provided includes truck ID and container number. Based on all the appointment requests associated with a truck ID, the TAS will determine the respective "desired" time-gaps between consecutive appointments. It is assumed that a truck will seek to perform as many container moves as possible in a given day, and thus, the time-gap between requested appointments represents the absolute minimum time the truck needs to complete the job; note that this time-gap includes in the truck turn time which is assumed to be constant. Once all of the input data (appointment requests and quotas) are provided, the TAS then determines the appointment time-window for each request with the objective of minimizing the cost of changing the time-gap between two consecutive appointments of a truck with more than one appointment, cost of changing an appointment to another time-window, and cost of queuing at the gate due to congestion (step 3). Next section provides the mathematical formulation of TAS. Lastly, the optimal appointments for each truck are sent to the drayage firms (step 4). The assigned appointment time-windows may require drayage firms to reschedule their trucks' tours if they are different from the requested times. The economic impact of TAS on drayage firms can be determined by computing the difference between the total drayage cost of the original tours and the cost of the adjusted tours.

When formulating the TAS, it is assumed that all trucks are required to make an appointment as it is in practice by Eagle Marine Services, eModal. It is also assumed that the container terminal will process trucks during the appointment time-window and that no appointment will be transferred to the next time-window because of the lack of resources at the container terminal. In practice, trucks are able to pick up and deliver containers before and after terminal hours; thus, it is assumed that a truck can pick up the container from the customer up to 4 hours before the terminal opens and deliver the container to the customer up to 4 hours after the terminal closes. We also assumed that drayage firms do not have knowledge about other requested appointments and the terminal-specified quotas.

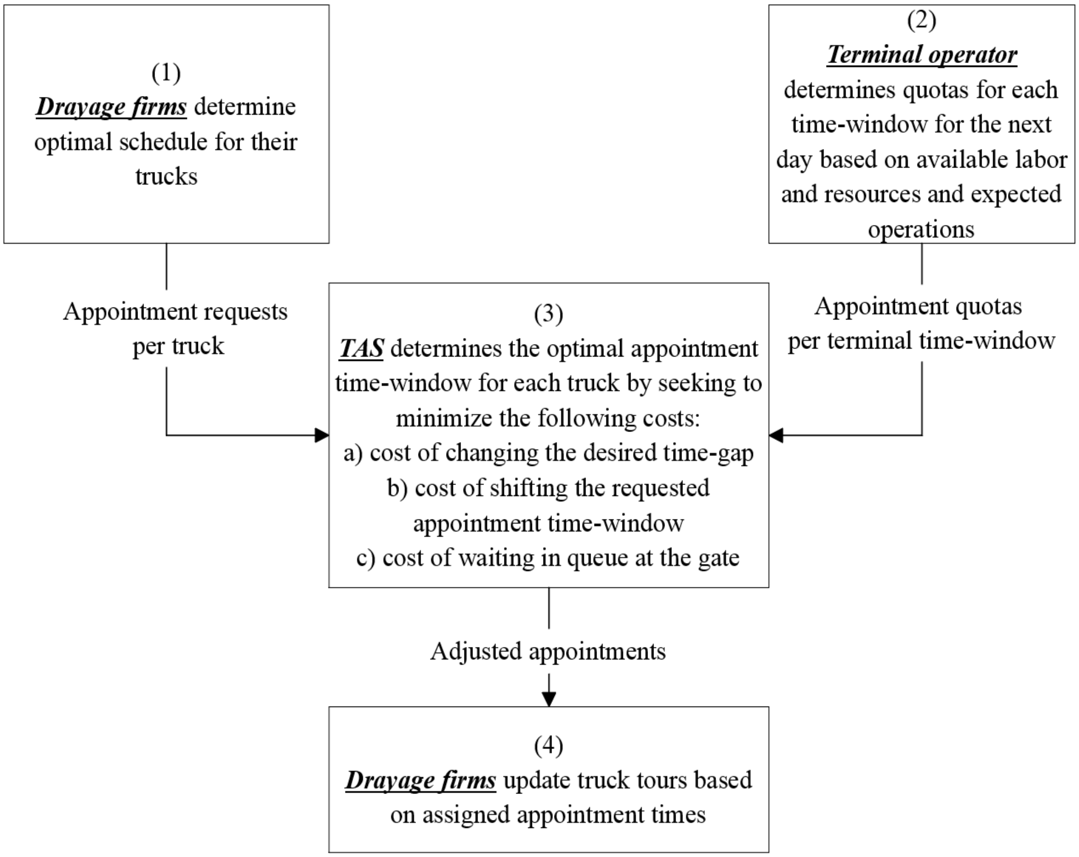


Figure 2.1 Illustration of the TAS framework

2.2.1 DRAYAGE SCHEDULING PROBLEM

Port drayage refers to the movement of containers between a maritime container terminal and an inland distribution point or rail terminal. A drayage job involves either delivering an outbound container to the maritime container terminal or picking up an inbound container from a maritime terminal.

The drayage problem in this study is adopted from the work of Shiri and Huynh (2016) which is a graph-based model. Nodes are used to represent inbound/outbound jobs which include several associated activities. The activities associated with an inbound job are terminal turn time, travel time between terminal and customer's location, container unmounting (i.e., unloading a container from a truck) time, container unpacking (i.e., stripping a container) time, and container mounting (loading a container onto a truck) time. Similarly, the activities associated with an outbound job are container unmounting time, container packing (i.e., stuffing a container) time, container mounting time, travel time between customer's location and terminal, gate queuing time, and terminal turn time. The arcs are used to represent the transfer time between the completion of first node activities and the commencement of the next node activities. The transfer time for all possible combinations are shown in Table 2.2.

Table 2.3 Transfer time between every two jobs (Shiri and Huynh, 2016)

From node	*To node*		
	Depot	*Inbound*	*Outbound*
Depot	NA	• Travel time between depot and terminal. • Gate queuing time.	• Travel time between truck depot and empty container depot. • Container mount time. • Travel time between empty container depot and customer.
Inbound	• Travel time between customer and empty container depot. • Container unmount time. • Travel time between empty container depot and truck depot.	• Travel time between first customer and empty container depot. • Container unmount time. • Travel time between empty container depot and terminal. • Gate queuing time.	• If customers are different, travel time between first customer to second customer should be considered. Otherwise, no activity.
outbound	• Travel time between terminal and truck depot.	• No activity.	• Travel time between terminal and empty container depot. • Container unmount time. • Travel time between empty container depot to customer.

2.2.2 TAS PROBLEM INDICES, PARAMETERS AND SETS

Indices

τ = Time-window

t = Time-interval

k = Truck ID

i = Number of appointments for a truck

a = Appointment number for a truck

d = Drayage firm ID

Sets

T = Set of time-windows, $T = \{1, 2, ..., 10\}$; e.g., T = 1 is the time-window from 8 AM to 9 AM

T_τ = Set of time-intervals within a time-window; e.g., $T_1 = \{1, 2, ..., 10\}$, $T_2 = \{1, 2, ..., 10\}$, etc.; $T_1 = 1$ is the time-interval from 8:00 AM to 8:06 AM

I = Number of possible appointments for each truck ranging from 1 to 5. Thus, the set I = {1, ..., 5}; e.g., I = {3} means that the truck has three appointments

R_i = Set of appointments for a truck with $i \in I$ appointment(s), $R_i = \{1, ..., i\}$; e.g., $R_i = \{1\}$, $R_i = \{1, 2\}$, etc.

D = Set of drayage firms, D = {1, 2, …}; e.g., D = 1 is the drayage firm 1

K_d = Set of trucks from drayage firm d; e.g., $K_1 = \{1, 2, ...\}$, $K_1 = 1$ is the first truck of drayage firm 1

Parameters

c_{τ} = Maximum number of appointments at time-window τ (quotas)

p_{kia} = Desired arrival time-window for ath appointment of truck k with i appointment(s)

μ_{t} = Maximum service rate (service capacity) of gate at time-window t

σ = Number of time-intervals in a terminal time-window

e = Coefficient of variance of gate service time

c^{gap}_{large} = Penalty value for actual time-gap larger than the desired time-gap

c^{gap}_{small} = Penalty value for actual time-gap smaller than the desired time-gap

c_{pos} = penalty value applied to truck that arrives earlier than scheduled

c_{neg} = penalty value applied to truck that arrives later than scheduled

c_{queue} = Penalty value for length of queue at the gate of terminal

n_{d} = Total number of appointment(s) requested from drayage firm d

TH_{d} = Threshold vale for the difference of cost between each drayage firm

$diffp_{kia}$ = Time-gap between ath and (a+1)th desired appointments of truck k with i appointments

2.2.3 TAS MATHEMATICAL FORMULATION

This work extends the work of Shiri and Huynh (2016) by (1) adding a new mathematical formulation described in section 3.3 for the TAS problem which is the main focus of this study, and (2) considering more than one drayage firm in the process of appointment reservation. In regard to contribution 1, the work of Shiri and Huynh (2016) is used to solve the *unrestricted* and *restricted* drayage problems (steps 1 and 4 in Figure

2.1), whereas the proposed new formulation is used to solve the TAS problem (step 3 in Figure 2.1). In regard to contribution 2, in practice, a typical maritime container terminal provides services to multiple drayage firms (Harrison et al. 2007; Phan and Kim, 2015 and 2016). The interplay between the proposed TAS model and the drayage scheduling model of Shiri and Huynh (2016) allows for the explicitly consideration of truck tours in optimizing truck appointments. Moreover, the proposed TAS model differs from the work of Phan and Kim (2015) in that it (1) considered a separate m-TSPTW model for drayage scheduling that return truck tours and their terminal arrival time-windows, (2) considered non-uniform quotas to account for terminal capacity fluctuations that may occur on a daily basis, and (3) ensured that the increase in tour cost for each individual drayage firm does not exceed a prespecified threshold; this threshold is different for each drayage firm.

Decision variable

Since the time-windows are discrete at terminal, a binary decision variable is defined to make the decision in the TAS model. The decision variable explains the arrival of a specific truck at a certain time-window at terminal.

$$x_{kia\tau} = \begin{cases} 1 & \text{If truck k with i appointment(s) has its } a^{th} \text{ appointment at time-window } \tau \\ 0 & \text{Otherwise} \end{cases}$$

Derived variables

λ_t = Average truck arrival rate at terminal gate at time-interval t

w_t = Average queue length at terminal gate at time-interval t

v_t = Average departure rate from terminal gate to the terminal yard at time-interval t

$diff_{kia}$ = Time-gap between a^{th} and $(a+1)^{th}$ appointments of truck k with i appointments

T_d = Total cost of appointment change for drayage firm d

N_{kia} = Difference between $diff_{kia}$ and $diffp_{kia}$

Q_{kia} = Difference between $diffp_{kia}$ and $diff_{kia}$

S_{kia} = Difference between actual and preferred arrival time-window

Z_{kia} = Difference between preferred and actual arrival time-window

$$\min \sum_{d\in D} T_d + c_{queue} \sum_{t\in T_\tau} (w_t + \frac{\lambda_t - v_t}{2}) \tag{2.1}$$

Subject to:

$$\begin{aligned} T_d = {} & c_{large}^{gap} \sum_{k\in K_d} \sum_{i\in I\setminus\{1\}} \sum_{a\in R_i} N_{kia} \\ & + c_{small}^{gap} \sum_{k\in K_d} \sum_{i\in I\setminus\{1\}} \sum_{a\in R_i} Q_{kia} \\ & + c_{pos} \sum_{k\in K_d} \sum_{i\in I} \sum_{a\in R_i} S_{kia} \\ & + c_{neg} \sum_{k\in K_d} \sum_{i\in I} \sum_{a\in R_i} Z_{kia} \end{aligned} \qquad \forall d \in D \tag{2.2}$$

$$N_{kia} \geq 0 \qquad \forall d \in D, \forall k \in K_d, \forall i \in I\setminus\{1\}, \forall a \in R_i \setminus \{i\} \tag{2.3}$$

$$N_{kia} \geq diff_{kia} - diffp_{kia} \qquad \forall d \in D, \forall k \in K_d, \forall i \in I\setminus\{1\}, \forall a \in R_i \setminus \{i\} \tag{2.4}$$

$$Q_{kia} \geq 0 \qquad \forall d \in D, \forall k \in K_d, \forall i \in I\setminus\{1\}, \forall a \in R_i \setminus \{i\} \tag{2.5}$$

$$Q_{kia} \geq diffp_{kia} - diff_{kia} \qquad \forall d \in D, \forall k \in K_d, \forall i \in I\setminus\{1\}, \forall a \in R_i \setminus \{i\} \tag{2.6}$$

$$S_{kia} \geq 0 \qquad \forall d \in D, \forall k \in K_d, \forall i \in I\setminus\{1\}, \forall a \in R_i \setminus \{i\} \tag{2.7}$$

$$S_{kia} \geq \sum_{\tau\in T} (\tau \cdot x_{kia\tau}) - p_{kia} \qquad \forall d \in D, \forall k \in K_d, \forall i \in I\setminus\{1\}, \forall a \in R_i \setminus \{i\} \tag{2.8}$$

$$Z_{kia} \geq 0 \qquad \forall d \in D, \forall k \in K_d, \forall i \in I \backslash \{1\}, \forall a \in R_i \backslash \{i\} \qquad (2.9)$$

$$Z_{kia} \geq p_{kia} - \sum_{\tau \in T} (\tau \cdot x_{kia\tau}) \qquad \forall d \in D, \forall k \in K_d, \forall i \in I \backslash \{1\}, \forall a \in R_i \backslash \{i\} \qquad (2.10)$$

$$\frac{T_d}{n_d} \leq TH_d \qquad \forall d \in D \qquad (2.11)$$

$$diff_{kia} = \sum_{\tau \in T} \tau \cdot x_{ki(a+1)\tau} - \sum_{\tau \in T} \tau \cdot x_{kia\tau} \qquad \forall d \in D, \forall k \in K_d, \forall i \in I \backslash \{1\}, \forall a \in R_i \backslash \{i\} \qquad (2.12)$$

$$diffp_{kia} = p_{ki(a+1)} - p_{kia} \qquad \forall d \in D, \forall k \in K_d, \forall i \in I \backslash \{1\}, \forall a \in R_i \backslash \{i\} \qquad (2.13)$$

$$\sum_{\tau \in T} x_{kia\tau} = 1 \qquad \forall d \in D, \forall k \in K_d, \forall i \in I, \forall a \in R_i \qquad (2.14)$$

$$\sum_{k \in K} \sum_{i \in I} \sum_{a \in R_i} x_{kia\tau} \leq c_\tau \qquad \forall \tau \in T \qquad (2.15)$$

$$\tau \cdot \sum_{\tau \in T} x_{kia\tau} \leq \tau \cdot \sum_{\tau \in T} x_{ki(a+1)\tau} \qquad \forall d \in D, \forall k \in K_d, \forall i \in I \backslash \{1\}, \forall a \in R_i \backslash \{i\} \qquad (2.16)$$

$$\lambda_t = \frac{\sum_{k \in K_d} \sum_{i \in I} \sum_{a \in R_i} x_{kia\tau}}{\sigma} \qquad \forall t \in T_\tau, \forall \tau \in T \qquad (2.17)$$

$$v_t \leq \mu_t \frac{w_t + 1 - \sqrt{w_t^{\,2} + 2e^2 \cdot w_t + 1}}{1 - e^2} \qquad \forall t \in T_\tau, \forall \tau \in T \qquad (2.18)$$

$$v_t \leq w_t + \lambda_t \qquad \forall t \in T_\tau, \forall \tau \in T \qquad (2.19)$$

$$w_{t+1} = w_t + \lambda_t - v_t \qquad \forall t \in T_\tau, \forall \tau \in T \qquad (2.20)$$

$$x_{kia\tau} = \{0 \text{ or } 1\} \qquad \forall d \in D, \forall k \in K_d, \forall i \in I, \forall a \in R_i \backslash \{i\}, \forall \tau \in T \qquad (2.21)$$

Eq. (2.1) is the objective function which seeks to minimize the total cost. The first term of the objective function is the total appointment change cost for all drayage firms. The last term represents the average waiting cost which is calculated from multiplying the average queue length between the beginning (w_t) and the end (w_{t+1}) of time-interval t by the cost of waiting for queue length (η). This concept is illustrated in Figure 2.2. Constraint (2.2) calculates the change in tour cost for every drayage firm. The first term of the equation is the cost of making the gap bigger than the desired gap for those trucks with more than one appointment (δ_{large}). The second term is the cost of making the gap smaller than the desired gap for those trucks with more than one appointment (δ_{small}). The third term is the cost of changing an appointment to a later time-window (γ_{pos}) and the fourth term is the cost of changing an appointment to an earlier time-window (γ_{neg}). Constraint (2.3) and constraint (2.4) are the linearization form of the maximum of zero and $diff_{kia} - diffp_{kia}$. Similarly, constraint (2.5) and constraint (2.6) are the linearization form of maximum of zero and $diffp_{kia} - diff_{kia}$. Constraint (2.7) and constraint (2.8) are the linearization form of maximum of zero and the difference between actual and preferred arrival time-windows. Constraint (2.9) and constraint (2.10) are the linearization form of maximum of zero and the difference between preferred and actual arrival time-windows. Constraint (2.11) ensures that the increase in tour cost for each individual drayage firm does not exceed a prespecified threshold. Instead of using a fixed threshold value for all drayage firms, the following function is proposed for determining the threshold value. The advantage of this approach is that it takes into account how many appointments are made by each drayage firm; a drayage firm with more appointment requests will have a lower threshold value.

$$TH_d = a + b \cdot h^{-n_d} \qquad \forall d \in D \tag{2.22}$$

Where parameter *a* represents the lowest threshold that the terminal operator wants to apply for all drayage firms. Parameter *b* represents the starting threshold for those drayage firms that have relatively few number of appointments. Parameter *h* represents the slope or rate of decreasing threshold.

Constraint (2.12) calculates the time-gap between two actual consecutive appointments of a truck with more than one appointment. Constraint (2.13) calculates the time-gap between two desired consecutive appointments of a truck with more than one appointment. Constraint (2.14) states that an appointment request should be met. Constraint (2.15) is the capacity constraint requiring the number of appointments in each time-window to be less than the specified quotas. The pre-specified quotas are calculated as follows:

$$AQ = \frac{\left(2\sum_{d \in D} n_d\right)}{10} \tag{2.23}$$

$$c_\tau = [1.1AQ, 1.1AQ, 1.1AQ, 0.9AQ, 0.9AQ, 0.9AQ, 1.1AQ, 1.1AQ, 1.1AQ, 1.1AQ] \tag{2.24}$$

Where AQ is average quota per time-window. It should be noted that all the experiments in this study are performed with 10 time-windows except for experiments related to sensitivity analysis of terminal time-window duration.

Constraint (2.16) ensures that the order of the adjusted appointments for a truck with multiple appointments is in the same order as requested. Constraint (2.17) adds up all of the truck arrivals in a time-window from different truck tours and divide them by the number of time-intervals in a time-window to calculate the average number of truck arrivals in a time-interval. In this study, the term time-interval is used to indicate a shorter duration than time-window; there are 10 time-intervals in a time-window. The use of time-interval is needed because the Pointwise Stationary Fluid Flow approximation (PSFFA) method used in constraints (2.17), (2.18), (2.19) and (2.20) requires a shorter duration than time-windows (Chen et al. (2011; 2013a; 2013b), Phan and Kim (2015)). Figure 2.2 shows the graphical representation of the PSFFA method. The term time-interval is only used for the purpose of queue length estimation. Constraint (2.18) and constraint (2.19) are to set the departure rate from the gate to the container yard at each time-interval to be the minimum of $\mu_t \frac{w_t+1-\sqrt{w_t^{\,2}+2e^2 \cdot w_t+1}}{1-e^2}$ and $w_t+\lambda_t$. It should be noted that it is not possible to linearize constraint (2.18) of the TAS model. The rest of the TAS model is linear. Constraint (2.20) calculates the truck queue length at each time-interval.

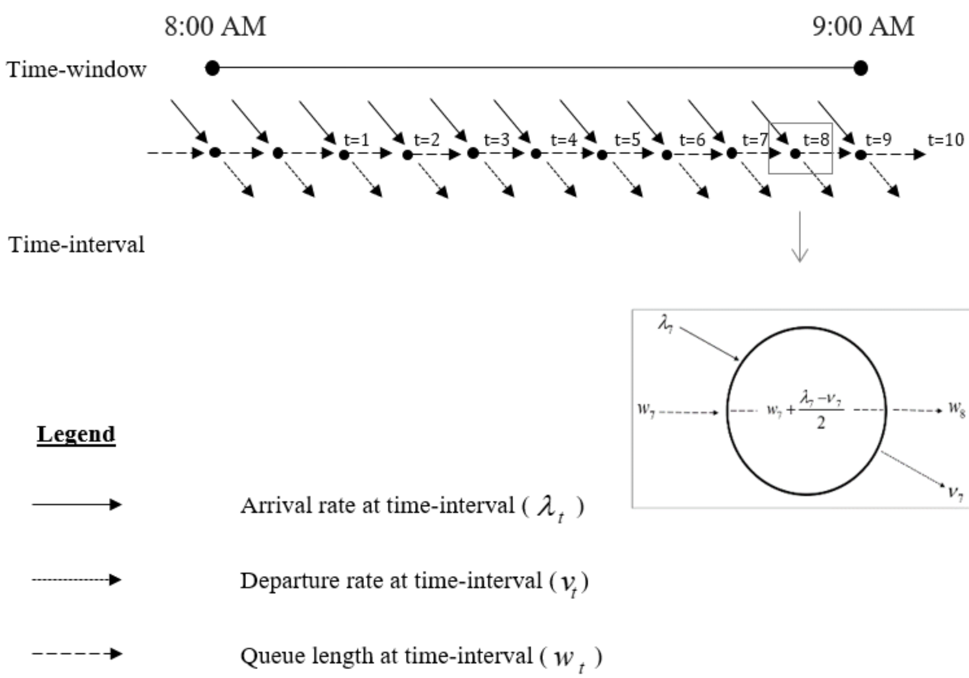


Figure 2.2 Illustration of PSFFA method used for queue length estimation

2.3 SOLUTION METHODOLOGY

The solution methodology is explained in context of the framework shown in Figure 2.1. In step 1, the drayage scheduling problem with no restriction on quota (referred to as the *unrestricted* drayage scheduling problem) is solved using CPLEX 12.6 (called from Python) to find the trucks' optimal tours for the scheduled jobs. Given that the drayage scheduling problem is NP-hard since it is an extension of the m-TSPTW (Shiri and Huynh, 2016), only small problems (up to 10 jobs per drayage firm) can be solved using CPLEX. But, the Reactive Tabu Search algorithm developed by Shiri and Huynh (2016) is used to deal with scenarios where a drayage firm need to process a larger number of jobs. The drayage firms will request the appointment time-window(s) that correspond to the desired truck arrival time(s) at the terminal. A Python program, as shown in Figure 2.3, is used in this study to determine the truck appointment time-window(s) from the truck tours. In step 2, the terminal operator specifies the quota per time-window. As done previously, we assume that the terminal appointment quota is 2 times higher than the total truck demand for the day (Shiri and Huynh, 2016; 2017). In step 3, given the truck appointment requests and quotas, the TAS model is solved such that the minimum cost of appointment change is achieved among drayage firms. To achieve this, the following substeps are performed. In step 3.1, set the initial parameters for the threshold equation (Eq. 2.22) as $a = 10, b = 4a, d = 1.25$ and solve the TAS model using LINGO 16.0 to obtain the optimal appointments. Note that LINGO is used to solve the TAS model instead of CPLEX because the TAS model is a nonlinear program; CPLEX can only solve linear, integer and mixed-integer linear programs. In step 3.2, calculate the percentage difference between the maximum and average T_d:

$$\text{TAS equality measure (EM)} = \frac{\underset{d\in D}{\text{Max}}\,(T_d) - \underset{d\in D}{\text{Average}}\,(T_d)}{\underset{d\in D}{\text{Average}}\,(T_d)} \times 100 \qquad (2.25)$$

In step 3.3, if the TAS model was infeasible, stop. Otherwise, lower the value of parameter *a* and go to step 3.1. Repeat these three steps until the TAS model becomes infeasible to solve using LINGO which means that the lowest TAS equality measure has been achieved.

In step 4, the adjusted appointment times are then sent to drayage firms. Lastly, from the given appointment times, each drayage firm will make adjustments to their trucks' tours; this is accomplished by solving the *restricted* drayage scheduling problem (i.e., drayage scheduling problem with quota restricted time-windows at container terminal).

2.4 NUMERICAL EXPERIMENTS

To demonstrate the feasibility of the developed model and solution methodology, they are applied on randomly generated experiments with real life characteristics. Experiments are generated on an instance square network of drayage area for the port of Los Angeles with deterministic travel times as shown in Figure 2.4. The use of the square network offers several benefits. First, it is easier to generate customer locations from the dimensions of a square. Second, any two customers generated for experiments are guaranteed to be within the square network. Lastly, it is often the case that the container terminal, empty container depot and chassis yard are located along the edge of the network. The network size is chosen such that a truck is able to make three turns per day on average. The size of the network and the location of the maritime container terminal are provided for every experiment. The location of the truck depot and empty container depot are randomly selected for different drayage firms. The customer locations are placed randomly within the network. The customers' pickup and delivery time-window is from 4:00 AM to

10:00 PM, and the container terminal operates from 8:00 AM to 6:00 PM. Ten terminal time-windows are used; thus, each time window is one-hour long.

```
for (d in drayage firms):
    tours = φ
    p_kia  = φ
    for (i in depots):
       for (j in customers):
           if (Y_ij = 1):
               Create a new tour and add customer j to the tour
               Create a new terminal arrival time-interval list
    untill (all customers are assigned to a tour):
        for (i in customers):
            for (j in depots and customers):
                if (Y_ij = 1):
                    for (k in tours):
                        if (last customer in tour k == customer i):
                               Add customer j to tour k
                               for (m in time-intervals):
                                      if q_im = 1:
                                         p_kia ← m
```

Figure 2.3 Pseudocodes of Python program used to determine trucks' preferred appointment times from unrestricted drayage model

Figure 2.4 Instance network of drayage area for the port of Los Angeles and Long Beach

The TAS model uses five different penalty parameters. To determine their values for our numerical experiments, we first performed an exhaustive enumeration of all possible combinations of values for these parameters. The values considered range from 1 to 10 (integers only). The optimal penalty values are those that yield the lowest difference between the *unrestricted* drayage scheduling model's objective function value and the *restricted* drayage scheduling model's objective function value. The optimal penalty values were found to be $c_{large}^{gap} = 1$ unit of cost/time-window, $c_{small}^{gap} = 3$ unit of cost/time-window, $c_{pos} = 1$ unit of cost/time-window, $c_{neg} = 3$ unit of cost/time-window, and $c_{queue} = 1$ unit of cost/length of queue. The unit of cost of these penalty values are relative to each other which suggest that the cost of shifting an appointment to a later time and the cost of assigning a time-gap between appointments smaller than the desired time-gap is the same. Their unit costs are 3 times higher than others. These penalty values are used for all experiments in this study. In the work by Phan and Kim (2015), the authors set the penalty values as follows: $c_{pos} = 1, c_{neg} = 4$, and $c_{queue} = 3$; a higher c_{queue} value suggests that the authors put more emphasis on minimizing gate congestion. It should be noted that different combinations of penalty values do not necessarily lead to a unique solution. That is, two different combinations of penalty values may produce the same truck appointments. All experiments are conducted on a desktop computer with Intel Core i7 3.4 GHz CPU and 8 GB of RAM.

2.4.1 EXPERIMENTAL DATA

The generated experiments used actual terminal data wherever possible or data documented in published manuscripts. An average turn time of 44.05 minutes is used for the experiments in this paper. It is based on the average turn time at the Port of Los Angeles and Long Beach where it was 45.2 minutes in November 2016 and 42.9 minutes in

December 2016 (PierPass). A combination of 71% of inbound full containers and 29% of outbound full containers are used for the experiments in this paper. It is based on the total volume of 6,363,250 TEUs at the port of Los Angeles and Long Beach in 2016 Calendar Year (CY) including 4,544,747 TEUs of inbound containers and 1,818,501 TEUs of outbound containers. The time to mount/unmount the container at customer location is assumed to be 5 minutes (Chung et al. 2007). Packing/unpacking times are assumed to be uniformly distributed with a minimum of 5 minutes and a maximum of 60 min, U (5, 60) (Zhang et al., 2010). The container terminal is assumed to operate 10 hours each day and there are a total of 10 time-windows (Shiri and Huynh, 2016). The average queuing time at the terminal is assumed to be 10 minutes (Chen et al., 2013c; Shiri and Huynh, 2016). The size of the drayage area and the location of the maritime container terminal represents the geographic layout of the Port of Los Angeles and Long Beach as shown in Figure 2.4; the travel time on each edge of the rectangle varies between 3 to 5 hours (Google Maps, 2017).

2.4.2 ILLUSTRATIVE EXAMPLES

Due to the complexity of the models, their inner workings are illustrated via several simple experiments. Table 2.3 provides a summary of the parameters used for the simple experiments. The results of the simple experiments are provided in Table 2.4. Row (1) lists the experiment number. The numbers in the square brackets in Row (2) indicates the arrival time-windows for a truck obtained from the unrestricted drayage scheduling model. Row (3) shows the objective function value of the unrestricted drayage scheduling problem. Row (4) provides the total number of adjusted arrivals per time-window obtained from the TAS model. Row (5) shows the adjusted arrival time-windows for every truck.

Row (6) lists the different cost components of the TAS objective function. Row (7) presents the new arrival time-windows for every truck obtained from the restricted drayage scheduling model. Row (8) provides the objective function value of the restricted drayage scheduling model.

Table 2.4 Parameters of simple experiments with one drayage firm

Experiment number	1	2	3
Quota per time-window	(1,1,0,1,1,1,1,1,0,1)	(1,1,1,1,1,1,1,1,1,1)	(2,2,2,2,2,2,2,2,2,2)
Problem size	4	5	5
Number of drayage firms	1	1	1
Number of inbound jobs	3	3	3
Number of outbound jobs	1	2	2
Terminal x-coordinate (min)	120	120	120
Terminal y-coordinate (min)	120	120	120
Drayage area (min x min)	120x120	120x120	120x120

Table 2.5 Results of simple experiments with one drayage firm

(1)	Experiment number		1	2	3
(2)	*Unrestricted* drayage firm 1 solution	Desired arrival time-windows for every truck	[1,3,6,8]	[2,2] [1,10,10]	[2,2] [1,10,10]
(3)		Objective function value	726.46	1755	1755
(4)	TAS solution	Adjusted arrivals per time-window	(1,0,0,1,0,1,0,1,0,0)	(1,1,1,0,0,0,0,0,1,1)	(1,2,0,0,0,0,0,0,0,2)
(5)		Adjusted arrival time-windows for every truck	[1,4,6,8]	[2,3] [1,9,10]	[2,2] [1,10,10]
(6)		Total objective value	Total = 10.2 η = 5.2 $\delta_{large} = 1$ $\delta_{small} = 3$ $\gamma_{pos} = 1$ $\gamma_{neg} = 0$	Total = 15.0 η = 6.0 $\delta_{large} = 2$ $\delta_{small} = 3$ $\gamma_{pos} = 1$ $\gamma_{neg} = 3$	Total = 6.0 η= 6.0 $\delta_{large} = 0$ $\delta_{small} = 0$ $\gamma_{pos} = 0$ $\gamma_{neg} = 0$
(7)	*Restricted* drayage firm 1 solution	New arrival time-windows for every truck	[4] [1,6,8]	[1] [9,10] [2,3]	[2,2] [1,10,10]
(8)		Objective value	793.63	2003	1755

Table 2.6 All feasible solutions of the TAS model for experiment 1

Solution number	Solution (arrival time-windows)	δ_{large} (unit cost)	δ_{small} (unit cost)	γ_{pos} (unit cost)	γ_{neg} (unit cost)	η (unit cost)	Objective function value (total cost)
1	**[1, 4, 6, 8]**	**1**	**3**	**1**	**0**	**5.2363**	**10.2363**
2	[2, 4, 6, 8]	0	3	2	0	5.2364	10.2364
3	[1, 4, 7, 10]	2	0	4	0	4.4271	10.4271
4	[2, 4, 7, 10]	1	0	5	0	4.4272	10.4272
5	[1, 4, 7, 8]	1	3	2	0	5.3683	11.3683
6	[2, 4, 7, 8]	0	3	3	0	5.3684	11.3684
7	[1, 4, 8, 10]	2	0	5	0	4.4271	11.4271
8	[2, 4, 8, 10]	1	0	6	0	4.4272	11.4272
9	[1, 2, 6, 8]	1	3	0	3	5.3684	12.3684
10	[1, 5, 8, 10]	2	0	6	0	4.4271	12.4271
11	[2, 5, 8, 10]	1	0	7	0	4.4271	12.4271
12	[1, 4, 6, 10]	3	3	3	0	4.4272	13.4272
13	[2, 4, 6, 10]	2	3	4	0	4.4273	13.4273
14	[1, 2, 5, 8]	1	3	0	6	5.3683	15.3683
15	[1, 5, 6, 8]	2	6	2	0	5.3686	15.3686
16	[2, 5, 6, 8]	1	6	3	0	5.3686	15.3686
17	[1, 5, 7, 10]	3	3	5	0	4.4272	15.4272
18	[2, 5, 7, 10]	2	3	6	0	4.4272	15.4272
19	[1, 2, 6, 10]	3	3	2	3	4.5593	15.5593
20	[1, 5, 7, 8]	2	6	3	0	5.3685	16.3685
21	[2, 5, 7, 8]	1	6	4	0	5.3685	16.3685
22	[1, 4, 6, 7]	1	6	1	3	5.3685	16.3685
23	[2, 4, 6, 7]	0	6	2	3	5.3685	16.3685
24	[4, 5, 8, 10]	0	3	9	0	4.5593	16.5593
25	[1, 2, 7, 10]	3	3	3	3	4.5593	16.5593
26	[1, 4, 5, 8]	2	6	1	3	5.3683	17.3683
27	[1, 2, 5, 7]	0	3	0	9	5.3684	17.3684
28	[2, 4, 5, 8]	1	6	2	3	5.3685	17.3685
29	[1, 6, 8, 10]	3	3	7	0	4.4272	17.4272

Table 2.7 All feasible solutions of the TAS model for experiment 1 (Continued)

Solution number	Solution (arrival time-windows)	δ_{large} (unit cost)	δ_{small} (unit cost)	γ_{pos} (unit cost)	γ_{neg} (unit cost)	η (unit cost)	Objective function value (total cost)
30	[2, 6, 8, 10]	2	3	8	0	4.4272	17.4272
31	[4, 6, 8, 10]	0	3	10	0	4.4273	17.4273
32	[1, 2, 7, 8]	2	6	1	3	5.5005	17.5005
33	[1, 2, 8, 10]	3	3	4	3	4.5593	17.5593
34	[1, 2, 6, 7]	1	6	0	6	5.5005	18.5005
35	[1, 5, 6, 10]	4	6	4	0	4.5593	18.5593
36	[1, 2, 5, 10]	3	3	2	6	4.5593	18.5593
37	[2, 5, 6, 10]	3	6	5	0	4.5593	18.5593
38	[1, 4, 5, 7]	1	6	1	6	5.3686	19.3686
39	[2, 4, 5, 7]	0	6	2	6	5.3687	19.3687
40	[4, 5, 6, 8]	0	9	5	0	5.5394	19.5394
41	[4, 5, 7, 10]	1	6	8	0	4.5595	19.5595
42	[4, 5, 7, 8]	0	9	6	0	5.5008	20.5008
43	[1, 4, 5, 10]	4	6	3	3	4.5593	20.5593
44	[1, 6, 7, 10]	4	6	6	0	4.5593	20.5593
45	[2, 6, 7, 10]	3	6	7	0	4.5593	20.5593
46	[2, 4, 5, 10]	3	6	4	3	4.5594	20.5594
47	[4, 6, 7, 10]	1	6	9	0	4.5594	20.5594
48	[1, 6, 7, 8]	3	9	4	0	5.5391	21.5391
49	[2, 6, 7, 8]	2	9	5	0	5.5391	21.5391
50	[1, 5, 6, 7]	2	9	2	3	5.5391	21.5391
51	[2, 5, 6, 7]	1	9	3	3	5.5391	21.5391
52	[4, 6, 7, 8]	0	9	7	0	5.5392	21.5392
53	[5, 6, 8, 10]	0	6	11	0	4.5595	21.5595
54	[1, 2, 4, 8]	2	6	0	9	5.3686	22.3686
55	[1, 7, 8, 10]	4	6	8	0	4.5593	22.5593
56	[2, 7, 8, 10]	3	6	9	0	4.5593	22.5593
57	[4, 7, 8, 10]	1	6	11	0	4.5593	22.5593
58	[5, 7, 8, 10]	0	6	12	0	4.5594	22.5594

Table 2.8 All feasible solutions of the TAS model for experiment 1 (Continued)

Solution number	Solution (arrival time-windows)	δ_{large} (unit cost)	δ_{small} (unit cost)	γ_{pos} (unit cost)	γ_{neg} (unit cost)	η (unit cost)	Objective function value (total cost)
59	[4, 5, 6, 10]	2	9	7	0	4.7301	22.7301
60	[1, 2, 5, 6]	0	6	0	12	5.5005	23.5005
61	[1, 2, 4, 7]	1	6	0	12	5.3686	24.3686
62	[5, 6, 7, 10]	1	9	10	0	4.7301	24.7301
63	[1, 4, 5, 6]	1	9	1	9	5.5391	25.5391
64	[2, 4, 5, 6]	0	9	2	9	5.5392	25.5392
65	[1, 2, 4, 10]	4	6	2	9	4.5595	25.5595
66	[5, 6, 7, 8]	0	12	8	0	5.7229	25.7229
67	[4, 5, 6, 7]	0	12	5	3	5.7229	25.7229
68	[1, 2, 4, 6]	0	6	0	15	5.3687	26.3687
69	[6, 7, 8, 10]	0	9	13	0	4.7301	26.7301
70	[1, 2, 4, 5]	0	9	0	18	5.5008	32.5008

Experiment 1 involves one drayage firm which has four trucks available and it has to process four jobs: three inbound jobs and one outbound job. The results of experiment 1 show that the total cost for the *unrestricted* drayage scheduling problem is 726.46. The optimal solution for the *unrestricted* drayage problem requires the use of only one truck and its desired arrival time at the terminal are during time-windows: 1, 3, 6, and 8. Table 2.5 lists all of the TAS feasible solutions for experiment 1, and it highlights the TAS solution with the minimum objective function value (i.e., solution 1 is the optimal solution provided by the TAS model). The solutions are reported with the cost components of the objective function. The TAS optimal solution indicates that the optimal appointment times are at time-windows: 1, 4, 6, and 8. Note that the desired appointment at time-window 3 has been shifted to time-window 4 because the quota for time-window 3 is zero. Given the assigned truck appointments from the TAS, the solution to the *restricted* drayage scheduling problem is shown in row (7) of Table 2.4. The total cost of the *restricted* drayage scheduling problem is 793.63. This result means that the appointment system increased the drayage cost for the firm by 9.25%. The result of the *restricted* drayage scheduling problem indicates that two trucks are required. One truck will serve one customer and it will arrive at the terminal at time-window 4. The other truck will serve three customers and it will arrive at the terminal at time-windows 1, 6, and 8, respectively.

To further illustrate the inner workings of the models, the solutions of two other experiments (will be referred to as experiments 2 and 3) are provided in Table 2.4. Experiments 2 and 3 examine the scenarios where the appointment quotas are less restrictive. For experiment 2, the quotas at all time-windows are 1. Note that there are two appointment requests during time-windows 2 and 10. It should be noted that drayage scheduling model considers two appointments for a truck with double moves (e.g., dropping off an outbound container at terminal and picking up an inbound container from terminal on the same trip). The TAS model shifts one of the arrivals from time-window 2 to time-window 3, and one of the arrivals at time-window 10 to time-window 9. With these assigned appointments, the TAS model yields the following adjusted arrivals per time-window (1,1,1,0,0,0,0,0,1,1). These adjusted arrivals are used as quotas to the *restricted* drayage scheduling model. Since the quotas are different than the desired arrival time-windows, the objective function of the *restricted* drayage scheduling model is higher than the *unrestricted* drayage model. In changing the appointment times, the TAS increased the total drayage cost for the firm by 14.13%.

In experiment 3, the number of appointment requests in every time-window is less than the quota. Given the available capacity, the TAS does not change any of the appointments. For this reason, the drayage firm does incur any additional drayage cost. It should be noted that the reason why the drayage cost remains the same is because the number of trucks needed is 2 for both the *unrestricted* and *restricted* case. In experiment 2, the drayage cost increased because the number of trucks went from 2 in the *unrestricted* case to 3 in the *restricted* case.

Three additional small instances and their corresponding optimal solutions are provided in the Appendix A.

2.4.3 NUMERICAL EXPERIMENTS AND RESULTS

Table 2.6 provides a summary of the parameters of 34 additional experiments. Column 1 shows the experiment number. Column 2 represents the size of the experiment in terms of number of jobs. Column 3 shows the average packing/unpacking time for all drayage firms. Column 4 indicates the appointment quotas for the TAS problem which can be described as step 2 of the solution methodology. Column 5 shows the number of drayage firms involved in the experiment. Columns 6, 7 and 8 indicate the parameters of the Eq. (2.22). Columns 9 and 10 show the number of inbound and outbound jobs in the experiment. Columns 11 and 12 provide the x and y coordinates of the terminal. Lastly, column 13 shows the drayage area.

The results of the numerical experiments are shown in Table 2.7. Column 1 shows the experiment number. Columns 2, 3 and 4 represent the objective value, CPU time and number of required trucks for the unrestricted drayage problem. It should be noted that columns 2, 3 and 4 are the results of step 1 of the solution methodology. Columns 5 and 6 show the objective function value and CPU time for the Queue-based TAS problem. These columns are the results of step 3 of the solution methodology. Columns 7, 8 and 9 represent the objective value, CPU time and number of required trucks for the restricted problem under Queue-based TAS; these are the results of step 4 of the solution methodology. A queue-based TAS is one that only takes the trucks' queuing time into account. The queue-based TAS is a simplified TAS model which is based on PSFFA method and similar to one proposed by Zhang et al. (2013), Chen et al. (2013) and Guan and Liu (2009a). To model the queue-based TAS, we used the proposed TAS and set the penalty values of c^{gap}_{large}, c^{gap}_{small}, c_{pos}, and c_{neg} to zero. The results indicate that the proposed TAS yields the lowest number of trucks required for a drayage firm. Column 10 and 11 indicate the objective value and CPU time of the proposed TAS. Column 12 to column 16 show the five different components of the objective value for the proposed TAS problem. Column 17 shows the TAS equality measure for the proposed TAS. It should be noted that column 10 to column

17 are the outcome of step 3 of the solution methodology. Columns 18, 19 and 20 represent the objective value, CPU time and number of required trucks for the restricted drayage problem under proposed TAS. The last three columns are the results of step 4 of the solution methodology. Column 21 shows the percentage difference (Gap) in objective value between Queue-based TAS and proposed TAS. The gap is calculated as 100*[(column 7 – column 18)/column 2].

It can be seen that the objective function values of the *unrestricted* drayage problem are always lower than the objective function values of the *restricted* drayage problem. The average difference in cost is 10.1% for small-sized experiments (experiments 4 to 24) and 12.8% for large-sized experiments (experiments 25 to 33) compared to a queue-based TAS.

Typically, drayage firms want to minimize the number of trucks needed to process the requested jobs. According to columns 4, 9, and 14 in Table 2.7, the number of trucks required for the *unrestricted* drayage problem, *restricted* drayage problem with a queue-based TAS, and *restricted* drayage problem with proposed TAS.

To demonstrate that our proposed model can handle real-world scenarios, an experiment was conducted (experiment 33) with 4,180 jobs. The computation time for this experiment for *unrestricted* drayage problem, proposed TAS, and *restricted* drayage problem were 26.0, 12.6, and 17.0 minutes, respectively. The proposed model can run larger-sized problems; however, the computation time will get significantly longer.

Table 2.9 Parameters of the TAS and drayge problems

(1) Exp Num	(2) Size (jobs)	(3) Ave. packing/unpacking time (min)	(4) Quota per time-window	(5) Num. of drayage firm(s)	(6) $TH_d = a + b \cdot h^{-n_d}$ a	(7) b	(8) d	(9) Number of jobs In*	(10) Out*	(11) Terminal coordinates (min) x	(12) y	(13) Drayage area (min x min)
4	3	19.0	(1,1,1,1,1,1,1,1,1,1)	1	0.5	2	1.25	2	1	0	90	180x 180
5	8	7.5	(2,2,2,2,2,2,2,2,2,2)	1	0.7	2.8	1.25	5	3	0	90	180x 180
6	10	18.7	(3,3,3,2,2,2,3,3,3,3)	2	1.1	4.4	1.25	6	4	0	90	180x 180
7	14	15.8	(4,4,4,3,3,3,4,4,4,4)	2	0.2	0.8	1.25	8	6	0	90	180x 180
8	18	18.1	(4,4,4,4,4,4,4,4,4,4)	3	1.4	5.6	1.25	11	7	0	90	180x 180
9	22	12.7	(5,5,5,4,4,4,5,5,5,5)	3	0.3	1.2	1.25	14	8	0	90	180x 180
10	25	15.1	(6,6,6,5,5,5,6,6,6,6)	4	0.8	3.2	1.25	16	9	0	90	180x 180
11	33	12.1	(8,8,8,6,6,6,8,8,8,8)	4	0.7	2.8	1.25	21	12	0	90	180x 180
12	35	13.0	(8,8,8,7,7,7,8,8,8,8)	5	0.8	3.2	1.25	22	13	0	90	180x 180
13	38	13.1	(9,9,9,7,7,7,9,9,9,9)	5	1.3	5.2	1.25	25	13	0	90	180x 180
14	50	11.4	(11,11,11,9,9,9,11,11,11,11)	6	0.7	2.8	1.25	34	16	0	90	180x 180
15	56	13.2	(13,13,13,11,11,11,13,13,13, 13)	7	0.7	2.8	1.25	38	18	0	90	180x 180

Table 2.10 Parameters of the TAS and drayge problems (Continued)

(1) Exp Num	(2) Size (jobs)	(3) Ave. packing/unpacking time (min)	(4) Quota per time-window	(5) Num. of drayage firm(s)	(6) $TH_d = a + b \cdot h^{-n_d}$ a	(7) b	(8) d	(9) Number of jobs In*	(10) Out*	(11) Terminal coordinates (min) x	(12) y	(13) Drayage area (min x min)
16	62	13.8	(14,14,14,12,12,12,14,14,14,14)	8	0.7	2.8	1.25	42	20	0	90	180x180
17	70	14.9	(16,16,16,13,13,13,16,16,16,16)	9	0.6	2.4	1.25	45	25	0	90	180x180
18	83	12.8	(19,19,19,15,15,15,19,19,19,19)	10	0.6	2.4	1.25	54	29	0	90	180x180
19	95	13.1	(21,21,21,18,18,18,21,21,21,21)	13	0.7	2.8	1.25	61	34	0	90	180x180
20	119	12.6	(27,27,27,22,22,22,27,27,27,27)	16	0.6	2.4	1.25	77	42	0	90	180x180
21	133	12.4	(30,30,30,24,24,24,30,30,30,30)	19	0.6	2.4	1.25	85	48	0	90	180x180
22	175	12.5	(39,39,39,32,32,32,39,39,39,39)	22	0.7	2.8	1.25	114	61	0	90	180x180
23	187	12.6	(42,42,42,34,34,34,42,42,42,42)	25	0.6	2.4	1.25	119	68	0	90	180x180
24	219	12.5	(49,49,49,40,40,40,49,49,49,49)	28	0.6	2.4	1.25	140	79	0	90	180x180
25	293	13.1	(65,65,65,53,53,53,65,65,65,65)	40	0.6	2.4	1.25	189	104	0	90	180x180
26	391	12.9	(87,87,87,71,71,71,87,87,87,87)	50	0.8	3.2	1.25	257	134	0	90	180x180
27	461	12.7	(102,102,102,83,83,83,102,102,102,102)	60	0.8	3.2	1.25	299	162	0	90	180x180

Table 2.11 Parameters of the TAS and drayge problems (Continued)

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Exp Num	Size (jobs)	Ave. packing/unpacking time (min)	Quota per time-window	Num. of drayage firm(s)	$TH_d = a + b \cdot h^{-n_d}$			Number of jobs		Terminal coordinates (min)		Drayage area (min x min)
					a	b	d	In*	Out*	x	y	
28	546	12.7	(121,121,121,99,99,99,121,121,121,121)	70	0.7	2.8	1.25	350	196	0	90	180x180
29	582	12.8	(129,129,129,105,105,105,129,129,129,129)	80	1	4	1.25	375	207	0	90	180x180
30	667	13.1	(147,147,147,121,121,121,147,147,147,147)	90	0.7	2.8	1.25	432	235	0	90	180x180
31	733	12.0	(162,162,162,132,132,132,162,162,162,162)	100	0.7	2.8	1.25	471	262	0	90	180x180
32	1495	12.5	(329,329,329,270,270,270,329,329,329,329)	200	1.5	6	1.25	956	539	0	90	180x180
33	4180	12.7	(919,919,919,752,752,752,919,919,919,919)	500	1.4	5.6	1.25	2903	1277	0	90	180x180
34	582	12.6	(129,129,129,105,105,105,129,129,129,129)	80	2.7	10.8	1.25	415	167	0	90	180x180
35	582	12.3	(129,129,129,105,105,105,129,129,129,129)	40	1.7	6.8	1.25	410	172	0	90	180x180
36	582	12.6	(129,129,129,105,105,105,129,129,129,129)	10	3.3	13.2	1.25	408	174	0	90	180x180
37	582	12.9	(129,129,129,105,105,105,129,129,129,129)	5	7.6	30.4	1.25	409	173	0	90	180x180

*In: Inbound; *Out: Outbound

Table 2.12 Results of the TAS and drayge problems

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)
	Unrestricted drayage problem			Queue-based TAS					Proposed TAS											
				TAS		*Restricted* drayage problem			TAS								*Restricted* drayage problem			
Exp. Num.	Obj. (min)	CPU (s)	Num. Of required trucks	Obj. (total cost)	CPU (s)	Obj. (min)	CPU (s)	Num. Of required trucks	Obj. (total cost)	CPU (s)	δ_{large} (unit cost)	δ_{small} (unit cost)	γ_{pos} (unit cost)	γ_{neg} (unit cost)	η (unit cost)	EM (%)	Obj. (min)	CPU (s)	Num. Of required trucks	Gap (%)
4	655	0.69	1	2.4	0.62	816	0.62	2	4.4	0.80	1	0	1	0	2.4	0	656	1.43	1	24.4
5	2290	1.77	3	6.5	99.94	2513	1.20	3	15.7	2.15	0	3	0	6	6.7	0	2290	1.77	3	9.7
6	3124	1.31	4	8.1	1.00	3676	1.28	6	27.5	1.31	1	0	0	18	8.5	5	3125	0.71	4	17.6
7	3286	3.39	4	11.7	99.91	3640	1.40	6	12.2	0.60	0	0	0	0	12.2	0	3286	0.84	4	10.8
8	4584	4.42	7	15.3	99.84	4959	3.39	9	46.1	0.59	0	6	0	24	16.1	18	4584	3.38	7	8.2

Table 2.13 Results of the TAS and drayge problems (Continued)

9	5863	7.27	8	18.9	100.14	6482	4.62	11	28.8	0.64	0	0	9	0	19.8	29	5863	3.59	8	10.6
10	6929	4.44	9	21.8	99.68	7449	1.73	11	46.3	0.82	0	0	0	24	22.3	40	6929	0.59	9	7.5
11	8845	32.61	11	29.4	0.93	9176	38.89	13	56.9	1.84	0	9	0	18	29.9	29	9016	18.43	13	1.8
12	9191	24.74	13	31.3	1.21	9707	27.06	15	65.4	1.30	2	6	4	21	32.4	31	9306	18.32	14	4.4
13	10465	70.10	13	34.3	19.29	11412	44.24	18	103.3	135.98	2	15	0	51	35.3	28	10706	17.49	14	6.7
14	14206	61.68	18	46.0	1.45	15160	23.14	22	86.5	10.79	1	9	0	30	46.5	30	14324	35.96	19	5.9
15	15482	114.74	20	51.9	76.43	16328	103.71	24	100.6	10.16	0	6	0	42	52.6	27	15542	52.11	20	5.1
16	17624	107.20	21	57.8	0.74	19649	31.70	32	122.2	50.64	4	21	0	39	58.2	12	17666	64.29	22	11.3
17	17688	124.45	23	65.7	1.23	19467	23.91	34	120.6	12.44	6	24	0	24	66.6	15	17846	36.11	24	9.2
18	22428	80.78	28	78.7	202.04	24158	27.64	39	124.2	141.74	5	18	4	18	79.2	43	22789	44.23	31	6.1
19	25230	138.77	32	91.1	177.23	28625	83.45	47	149.2	1.09	0	12	1	45	91.2	67	25495	69.24	33	12.4
20	30979	73.77	41	114.6	9.69	35392	27.16	60	199.9	214.53	10	30	0	45	114.9	27	31305	42.77	43	13.2

Table 2.14 Results of the TAS and drayge problems (Continued)

21	34870	203.45	47	128.5	19.66	38744	94.63	64	191.6	99.55	6	21	0	36	128.6	71	35026	73.84	47	10.7
22	45861	169.02	59	170.6	537.65	52178	19.07	97	263.4	189.73	0	21	0	72	170.4	72	46122	127.92	61	13.2
23	50392	291.06	63	182.6	99.94	57335	302.94	96	287.4	193.92	12	36	0	57	182.4	50	51116	158.89	68	12.3
24	56062	306.98	75	214.5	111.00	63340	77.31	111	342.3	327.57	11	39	0	78	214.3	42	56767	192.20	80	11.7
Avg.	**18384**	**86.79**	**24**	**65.8**	**83.79**	**20486**	**44.72**	**34**	**114.0**	**66.58**	**3**	**13**	**1**	**31**	**66.2**	**30**	**18560**	**45.91**	**25**	**10.1**
25	77600	397.25	103	288.2	2113.20	84889	241.06	142	487.3	335.03	19	54	0	126	288.0	34	78098	140.64	108	8.8
26	104825	705.58	139	386.2	6120.77	114956	216.19	184	598.2	12.72	2	39	0	171	386.2	81	105064	301.23	140	9.4
27	121513	712.88	160	456.4	6015.11	135694	106.85	232	681.2	1693.04	0	36	0	189	456.2	82	121972	402.87	161	11.3
28	146026	769.83	186	549.1	15.40	172388	252.60	189	781.1	15.03	0	51	0	189	541.1	90	146294	336.77	189	17.9
29	154765	758.30	199	577.4	1504.62	175072	83.88	302	800.1	109.04	0	42	1	180	577.1	125	155445	403.42	203	12.7

Table 2.15 Results of the TAS and drayge problems (Continued)

30	172451	902.23	226	662.4	3667.33	193926	142.74	336	1028.1	27.76	0	75	0	291	662.1	76	173083	550.72	230	12.1
31	193233	1249.53	252	741.9	58.47	215296	521.29	258	1133.1	24.01	0	51	0	354	728.1	76	193808	720.06	254	11.1
32	394787	2082.63	514	1513.8	234.75	461714	843.75	521	2230.0	292.74	8	96	0	636	1490.0	140	395662	843.56	521	16.7
33	1114361	1560.24	1657	4428.2	355.88	1874589	757.71	1783	6589.1	408.33	0	357	0	1716	4108.1	120	1702464	1020.93	1703	15.4
Avg.	**275507**	**1015.39**	**382**	**1067.1**	**2231.73**	**380947**	**351.79**	**439**	**1592.0**	**324.19**	**3**	**89**	**0**	**428**	**1026.3**	**92**	**341321**	**524.47**	**390**	**12.8**

Table 2.16 Results of the drayge problems using Tabu Search and CPLEX

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)
	Unrestricted drayage problem						Proposed TAS								*Restricted* drayage problem					
	Tabu Search			CPLEX											Tabu Search			CPLEX		
Exp. Num.	Obj. (min)	CPU (s)	N.R.T*	Obj. (min)	CPU (s)	N.R.T*	Obj. (total cost)	CPU (s)	δ_{large} (unit cost)	δ_{small} (unit cost)	γ_{pos} (unit cost)	γ_{neg} (unit cost)	η (unit cost)	EM (%)	Obj. (min)	CPU (s)	N.R.T*	Obj. (min)	CPU (s)	N.R.T*
34	232486	7570.23	374	225092	86.28	259	1247	108.67	1	51	0	618	577	84.3	248554	8958.75	492	238620	54.63	267
35	225457	8009.24	320	NA	NA	NA	980	105.22	21	0	243	138	578	78.3	245793	7950.54	463	NA	NA	NA
36	251824	10971.23	294	NA	NA	NA	796	101.30	9	0	179	30	578	37.9	280506	11725.09	422	NA	NA	NA
37	265274	12581.59	310	NA	NA	NA	790	17.56	0	0	211	0	578	16.9	301241	12433.79	456	NA	NA	NA
Avg.	**243760**	**9783.07**	**325**	**225092**	**86.28**	**259**	**953**	**83.19**	**8**	**13**	**158**	**197**	**578**	**54**	**269024**	**10267.04**	**458**	**238620**	**54.63**	**267**

*N.R.T = Number of required truck(s)

As expected, in most of the experiments the δ_{small} cost is higher than the δ_{large} cost because δ_{small} has a higher penalty value than δ_{large}. Similarly, the γ_{neg} cost is higher than the γ_{pos} cost because γ_{neg} has a higher penalty value than γ_{pos}. The η cost at optimality is much higher than the other four costs because the penalty value of η is applied to all trucks in the queue at the gate for all time-windows; the other four penalty values are applied only when the requested appointment time-window is changed to a different one.

As mentioned in the solution methodology section, the Reactive-Tabu Search (RTS) algorithm developed by Shiri and Huynh (2016) can solve larger-sized problems. Table 2.8 shows the results of four experiments with the number of jobs 582 with number of drayage firms between 5 and 80 obtained using RTS. Experiment 34 with 80 drayage firms has an average of 7 jobs per drayage firm and this problem can also be solved by CPLEX. As shown in Table 2.8, maximum calculated gap between the RTS and CPLEX objective function value is (248554-238620) / 238620 * 100 = 4%. For Experiments 35, 36, and 37, CPLEX is unable to obtain the solution due to "out of memory" error.

Figure 2.5 shows how the ratio of inbound and outbound jobs affect the increase in drayage operation cost (difference between restricted and unrestricted). The results in Figure 2.5 indicate that when the inbound:outbound job ratio is between 40:60 and 80:20, the increase in cost is highest. The reason is because when the number of inbound and outbound jobs are out of balance, there are fewer opportunities for truckers to do double moves, where a trucker drops off an export container at the terminal and picks up an import on the same trip. Drayage firms typically look to maximize the number of double moves in their appointment scheduling. Therefore, any changes to the desired appointments imposed by the TAS will negatively affect the drayage firm.

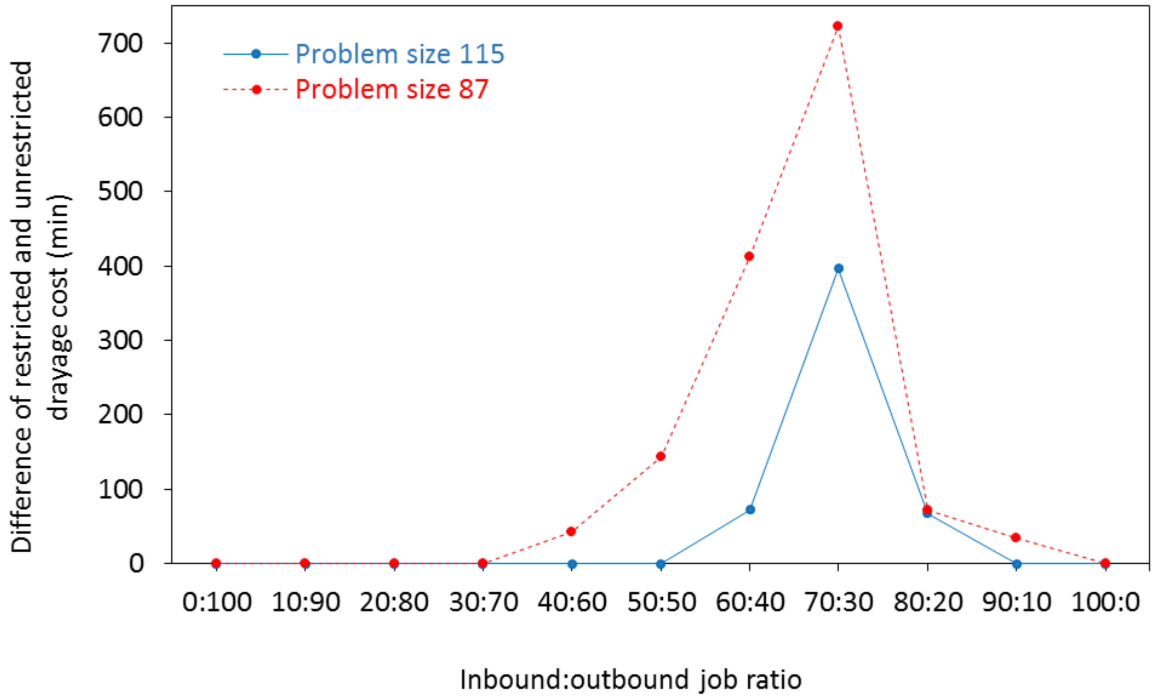


Figure 2.5 Effect of inbound:outbound job ratio on drayage cost

To test the effect of customer time-windows on drayage cost, five experiments with different customer time-windows are examined. Figure 2.6 shows the results of these experiments. As shown in Figure 2.6, the cost increases as the time-window decreases. Furthermore, the rate of cost increase is higher as the customer time-window becomes smaller.

To examine the effect of different parameters of the threshold equation, twelve different experiments are tested. Every four experiments are tested for a parameters while other two parameters are constant. The results of these experiments are plotted in Figure 2.7. It can be seen from the plot that the lowest TAS equality measure is obtained by decreasing parameters *a* and *b* and increasing parameter *h*.

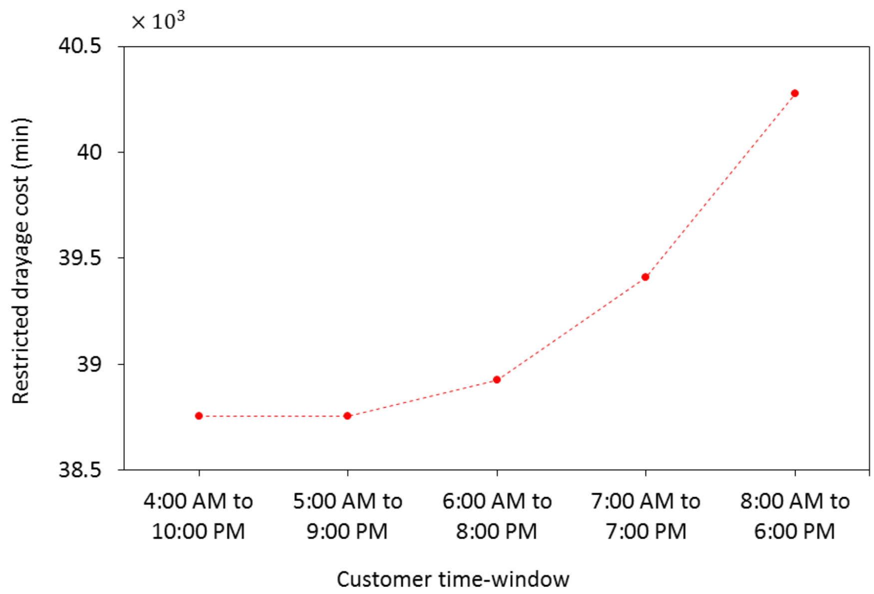


Figure 2.6 Effect of customer time-window on restricted drayage cost

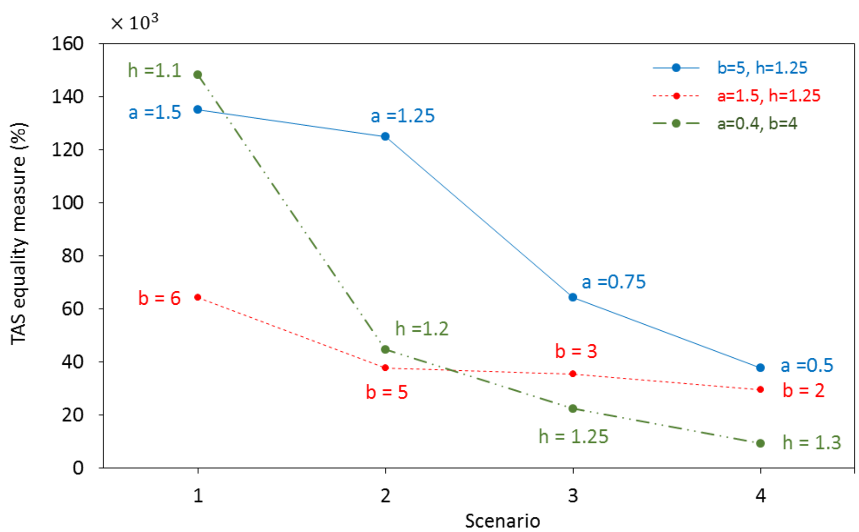


Figure 2.7 Effect of parameters of threshold equation on the TAS equality measure

The effect of appointment quotas is also examined. The ratio of quotas to requested appointments are varied between 3:1 to 0.8:1. As shown by the blue line in Figure 2.8, as the ratio of quotas to requested appointments decreases from 3:1 to 1.5:1, the restricted drayage cost increases slightly in a linear manner. The restricted drayage cost increases significantly when the ratio is 1:1. When the ratios are 0.9:1 and 0.8:1, the restricted drayage costs are significantly lower because some of the requested appointments cannot be accommodated by the terminal. The red line shown in Figure 2.8 represents the number of unsatisfied jobs. As shown, the number of satisfied jobs is zero when the ratios range between 3:1 and 1:1. At 0.9:1, the number of unsatisfied jobs is 9 and at 0.8:1, the number of unsatisfied jobs is 17.

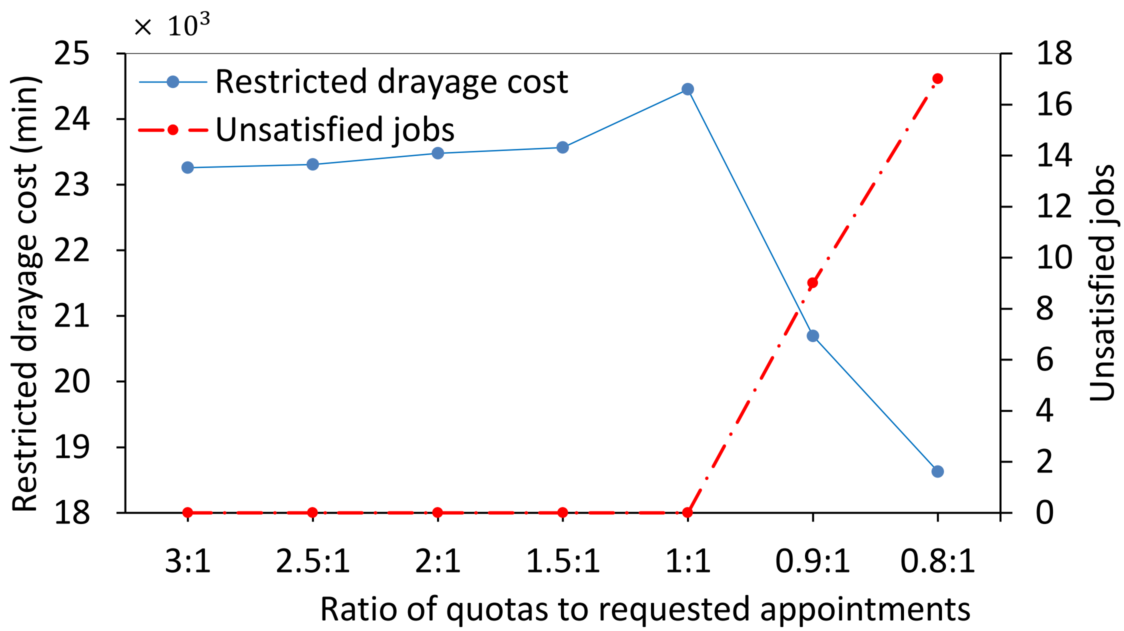


Figure 2.8 Effect of the ratio of quotas to requested appointments on drayage cost

The effect of terminal time-window duration is tested by changing it from 30 minutes to 300 minutes. The result of the analysis is plotted in Figure 2.9. As shown in Figure 2.9, increasing terminal time-window duration results in a reduction in drayage operation cost. As we expected, this reduction has a higher rate at small time-window periods than large time-windows. It should be noted that setting a narrow time-window could be beneficial to terminal operators because they can have a higher control on truck arrivals in every hour. But, wider time-windows results in lower drayage cost. Moreover, the arrival of trucks to the terminal is a function of highway congestion. Considering day-to-day travel time variation, a wider terminal time-window can increase the probability of being on-time for appointment.

The effect of terminal turn time on drayage cost is examined by 5 different experiments. Figure 2.10 shows the results of these experiments. As can be seen, terminal turn time is varied from 20 to 60 minutes on horizontal axis. The vertical axis shows the objective value of the drayage model for unrestricted and restricted drayage models. It is interesting to see that the cost of unrestricted and restricted drayage models become closer to each other as terminal turn time increases. It should be noted that increasing terminal turn time results in a lower number of jobs a truck can perform during a day and a higher drayage cost accordingly. For example, 60 minutes terminal turn time results in the highest drayage cost, so any changes in the requested appointments cannot result in a higher drayage cost.

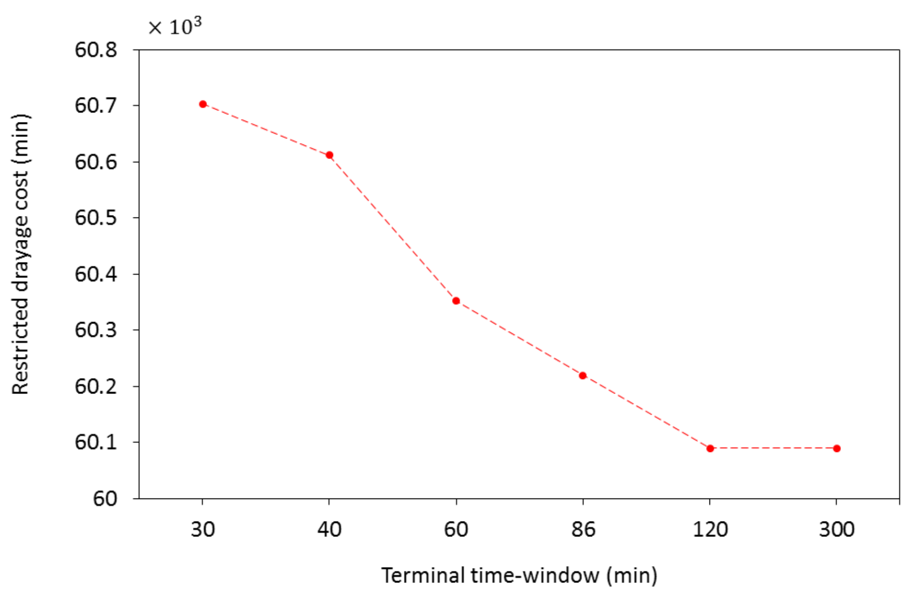


Figure 2.9 Effect of the size of terminal time-window on drayage cost

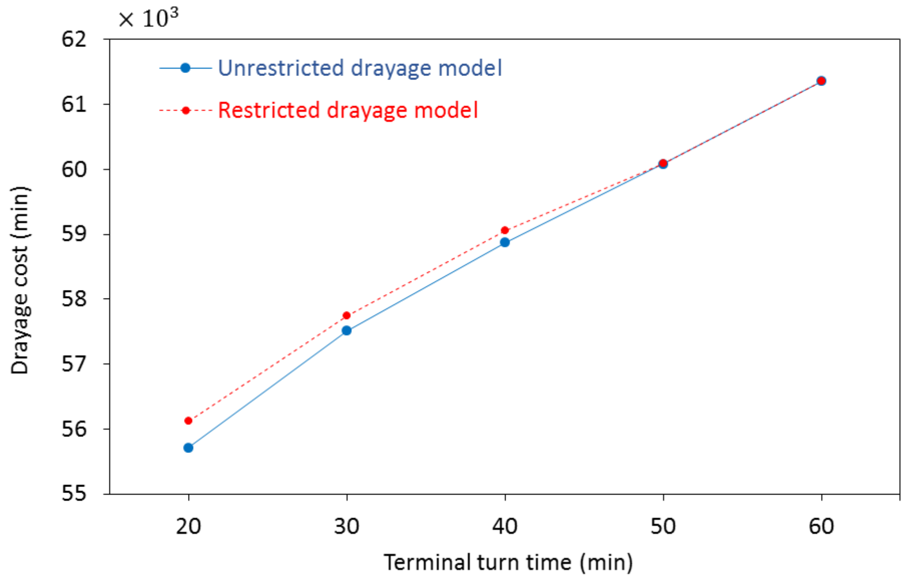


Figure 2.10 Effect of the terminal turn time on drayage cost

2.5 MANAGERIAL IMPLICATIONS

From a managerial point of view, our research has the following importance and key features that makes the proposed TAS design possible to be put into practice:

- *Incentive for high-volume drayage firms:* Using the proposed TAS, terminal operators are able to treat drayage firms based on their business volume (number of appointments requested). This enables port authorities to keep their top customers (drayage firms) satisfied with approving higher portion of their requested appointments without an adjustment. Using three parameters of the threshold equation (a, b, and h), port authorities can customize their approach to dealing with drayage firms of different volume.
- *Terminal resource utilization impact on drayage cost:* Using the sensitivity analysis performed on appointment quotas in every terminal time-window, the sensitivity analysis on terminal time-window duration, and the sensitivity analysis on terminal turn time, terminal operators can understand the impact of these changes on drayage cost.

This gives an insight to terminal operators that some of their above mentioned decisions may result in losing its business from some drayage firms because of an increase in their drayage cost.

- *Customer time-window impact on drayage cost:* Drayage firms' managers can understand the impact of customers' time-window on total drayage cost. The sensitivity analysis performed on customer time-window gives them an insight that having a customer with wider time-window results in a lower drayage cost while using the TAS.
- *Practicality:* First of all, the existing setup of drayage firms and maritime container terminals is such that a drayage firm does not share its customers' information such as truck scheduling, customers' request type (inbound/outbound), or customers' time-window with other drayage firms, maritime container terminal operators, nor the TAS. On the other side, terminal operators do not share terminal information such as expected turn time, resource utilization, or accurate container arrivals to the terminal with drayage firms. According to above mentioned restrictions on sharing information, our proposed TAS is compatible with this setting such that the TAS only uses the submitted appointment requests with a specific truck ID for an available truck in the terminal. Secondly, the proposed TAS requires only one time appointment request from drayage firms and one time appointment confirmation/adjustment from the TAS. So, it makes it very simple for both actors to be involved in this process. Finally, the computation time of the propose TAS for large size experiments is less than few minutes, using a personal computer, that makes it easy to be put into practice.

2.6 SUMMARY AND CONCLUSION

This paper addressed the challenge of designing a TAS that considers drayage firms' operation time as well as the maritime container terminal gate queuing time. First, the *unrestricted* drayage scheduling problem is solved to find the trucks' optimal tours for the scheduled jobs. In step 2, the terminal operator specifies the quota per time-window. Given the truck appointment requests and quotas, the TAS model is solved to obtain the optimal appointments (step 3). The adjusted appointment times are then sent to drayage firms. Lastly, from the given appointment times, each drayage firm will make adjustments to their trucks' tours; this is accomplished by solving the *restricted* drayage scheduling problem (step 4). It is found that considering truck tours in the TAS would decrease the operation cost of the *restricted* drayage scheduling model as well as the number of required trucks. Finally, it is found that when the inbound:outbound job ratio is between 40:60 and 80:20, the increase in drayage cost is highest.

To our knowledge, none of the existing truck appointment systems in the U.S. have all the critical features needed for long-term success and efficacy as discussed in the work of Huynh et al. (2016). This paper contributes to the existing body of work by proposing a centralized truck appointment system that considers both the interest of the terminal operator and drayage companies; it is the first to explicitly consider truck tours and seeks to provide appointment times such that trucks do not have to deviate greatly from their original schedule associated. Furthermore, it could be used to answer various system design issues that have yet to be examined. Such system design issues include: 1) what are the optimal values for appointment window duration, grace period, and appointment lead time? 2) what is the cost to the trucking industry as a result of mandatory appointment systems?; and 3) how should fees and penalties be structured to obtain optimal compliance from both drayage companies and terminal operators? With the proposed model and framework, decision makers could use it to examine how the appointment system parameters affect their respective operations as well as the system as a whole.

The limitations of this study need to be taken into account when interpreting the results. First, the hypothetical network used in the numerical experiments does not consider traffic congestion. Second, the distance between two locations is measured based on the Euclidean distance between two points. Lastly, the truck turn time at the terminal is assumed to be a constant across all time-windows. In future work, the research team plans to overcome some of these limitations by using an actual transportation network and by considering congestion on the highway network. In addition to relaxing the aforementioned limitations, another possible future research direction would be to formulate this problem as a bi-level mathematical program where the upper level seeks to determine the optimal appointment windows for trucks and the lower level seeks to determine the optimal routes for trucks given the assigned appointment windows

CHAPTER 3

MODELING THE TRUCK APPOINTMENT SYSTEM AS A MULTI-PLAYER GAME[1]

Marine container terminals play an important role in global trade. They provide the necessary land-sea connection for the shipments. The efficiency of marine container terminals affects supply chains and logistics of businesses. For example, the inability of the drayage operator to pick up an import container at the desired time may affect the on-time delivery of shipments to a retailer. A common occurrence at many large container terminals in the U.S. is gate congestion where long lines of trucks are waiting to enter the terminal. The extended queuing time increase the truck's turn time at the terminal and overall drayage operation time (Naboothiri and Erera, 2008). The increased drayage operation time reduces the drivers' available time to perform other moves, which in turn requires drayage firms to deploy more trucks to fulfill their pickup and delivery orders. More trucks on the road will further exacerbate traffic congestion, roadway safety, and even more congestion at terminals. Truck queuing at marine container terminals has serious public health implications. An idling truck emits a higher amount of emission than a moving truck (Brodrick et al., 2002; Moaris and Lord, 2006), and diesel truck emissions have been shown to trigger elevated levels of asthma attacks, emergency rooms visits,

[1]Torkjazi, Mohammad, Nathan Huynh, and Ali Asadabadi. "Modeling the Truck Appointment System as a Multi-Player Game."

hospitalizations, heart attacks, strokes and untimely deaths (Hill, 2005; Saxe and Larsen, 2004; Giuliano and O'Brien, 2007; Schulte et al., 2015, 2017).

Truck appointment system (TAS) or Vehicle Booking System (VBS) is one of the methods being used by terminal operators worldwide to reduce gate congestion (e.g., the Port of Baltimore, Port of Vancouver, and Port of Hamburg) (Heilig and Voß, 2017). The reasons and advantages of utilizing a TAS for the terminal operator and truckers are 1) the TAS will level out truck arrivals during a day and thereby reduce truck queueing time, 2) the TAS requires drayage firms to provide truck, chassis and container information prior to truck arrivals and thereby reduce gate processing time, 3) the TAS allows the terminal operator to set quotas during a day according to vessel operations, other workloads and available resources, and 4) the TAS provides a guarantee that the truck will be served when it arrives during the designated time-window.

In a typical TAS, terminal operators assign quotas to limit the number of trucks that can enter a certain yard area during a specific time-window. The quotas are typically set according to available resources. The time-window duration varies from terminal to terminal; they are generally between 1 to 4 hours. Most of the TAS pioneering studies aimed to determine the optimal quotas per time-window to reduce gate congestion (e.g., Huynh and Walton (2008), Huynh (2009); Chen et al. (2011); Chen et al. (2013a); Chen et al. (2013b); Zhang et al. (2013)). From the drayage firms' perspective, they first need to determine the optimal schedule (i.e., tour) for each of their trucks considering the number of moves and associated timing constraints. Then they would use the TAS to book appointments at time-windows that match their schedules. If the number of appointments made for a time-window exceed the specified quota, then they will need to look for another time-window. Requiring a truck to arrive at time-windows different from the ones desired effectively lowers drayage firms' productivity. Therefore, an effective TAS should consider not only gate congestion but also its effect on drayage firms, specifically, truck schedules. Recent TAS studies have begun to address this issue (e.g., Phan and Kim, 2015; 2016; Torkjazi et al., 2018).

The majority of TAS studies have developed models from the perspective of a single decision-maker that seeks to meet the interests of both the terminal operator and drayage firms (e.g., Namboothiri and Erera (2008); Phan and Kim, 2015; Torkjazi et al., (2018)); this type of models is referred to as single-player models herein. Single-player models require detailed operational information from all participating entities. In reality, the terminal, drayage firms and independent owner operators are separate entities that have different business interests and needs, and thus, single-player models may not be practical to implement. For this reason, recent studies have explored models, where each entity is a separate decision-maker with a separate objective (e.g., Chen et al., 2013; Phan and Kim, 2015; 2016); this type of models is referred to as multi-player models herein. The challenge with these models is that to obtain the equilibrium solution; the different models representing different entities need to be solved multiple times to capture the interplay between the terminal and drayage firms. This study seeks to overcome this limitation as explained below.

Given that the TAS problem is multi-player in nature with multiple rational decision makers, game theory is best suited for modeling this problem. Specifically, the TAS problem is treated as a multi-player game where each player seeks to optimize his objective function. It should be noted that the objectives of the players in this problem are conflicting. In this study, it is assumed that the players are not cooperating. To model this problem, a multi-player bi-level model is proposed, where the terminal functions as the leader at the upper-level, and drayage firms function as followers in the lower-level. The objective of the leader (the terminal) is to minimize the queuing time of trucks by spreading out the truck arrivals, and the objective of the followers (drayage firms) is to minimize their own drayage cost. The bi-level model is a mixed-integer nonlinear programming problem (MINLBP) in which the upper-level problem corresponding to the leader's objective is formulated as a mixed-integer program (MIP), and the lower-level problem corresponding to the drayage firms is formulated as a nonlinear problem (NLP). To obtain an exact solution, lemmas and linearization techniques are used to transform the bi-level model into a single-level MIP with continuous variables. To the authors' knowledge, this is the first study to model the TAS problem using game theory, and it is the first to propose the use of KKT conditions for the lower-level problem to enable the TAS problem to be solved as a standard MIP.

3.1 LITERATURE REVIEW

A number of studies have sought to model the TAS with the objective of minimizing the truck queuing time at the marine container terminal and difference between the assigned appointments and requested appointments. There are also studies that have sought to assess the effectiveness of different TAS designs. A review of the TAS state-of-the-art and the state-of-the-practice can be found in the work of Huynh et al. (2016). A summary of previous TAS studies is provided in Table 3.1. The following review focuses specifically on those studies that model the TAS problem as either single-player or multi-player. Here, the term "single-player model" refers to models with one central authority who sought to design the TAS to maximize benefits for *both* the terminal and drayage firms, whereas "multi-player model" refers to models that sought to do the same but treat the terminal and drayage firms as non-cooperating agents with conflicting objectives.

The single-player TAS model was first proposed by Chen et al. (2011). They proposed a nonlinear programming model to minimize the change in preferred appointments and truck turn time including terminal gate queuing time. They used a two-phase optimization approach to first find the optimal truck arrivals pattern and then find a pattern of time-varying tolls that results in optimal arrival pattern. They found that their proposed model allowed the terminal operators to fully utilize the terminal capacity, without significant loss in level-of-service. Phan and Kim (2015) sought to adjust truck arrivals so that terminal gate queuing time and inconvenience of trucks from changing their arrival times are minimized. They assumed in the CDM that the terminal has a dominating bargaining power over drayage firms. They used the CDM's solution as a reference point for comparison to their proposed multi-player TAS model, which will be discussed later in this section. Schulte et al. (2015, 2017) proposed a collaborative TAS between drayage firms and the terminal to reduce drayage costs and emissions. In their earlier work, they developed a discrete-event simulation model to assess coordinated truck appointments in a practical case of drayage. They found that the approach effectively reduces port-related truck emissions caused by avoidable empty trips. In their latter work, they proposed an optimization model to introduce a collaborative planning model to be operated within a TAS and to investigate its impact on emission and drayage cost objectives. They found that their model provides appropriately coordinated truck schedules and reduces truck emissions and cost. Shiri and Huynh (2016) proposed a single-player TAS model that considered multiple drayage firms that operate at a container terminal. They proposed a model to optimize the schedule of each truck with explicit consideration of terminal-specified quotas for appointments. They developed a mixed-integer programming model to solve the empty container allocation problem, vehicle routing problem and appointment booking problem in a single-player manner. To solve the integrated model, they proposed a combination of reactive Tabu Search and greedy algorithm. The authors found that the drayage operation time increases as terminal quotas decrease and a TAS that minimizes gate queuing time would benefit drayage firms considerably. Torkjazi et al. (2018) proposed a single-player TAS to minimize the impact to both terminal and drayage operations. Their proposed TAS distributes the truck arrivals evenly throughout the day to reduce gate congestion while considering drayage truck tours to avoid assigning appointment times that are significantly before or after their requested appointment times. They formulated the TAS as a mixed integer nonlinear program (MINLP) and found that the proposed TAS reduces drayage operation cost by 11.5% compared to a TAS whose sole objective is to minimize the gate queuing time. Jovanovic (2018) proposed a TAS to maximize the number of dray operations per day while considering number of appointments per truck, truckers' working hours, and travel distance of every dray. They found that the proposed TAS improved gate queuing time and drivers' satisfaction.

A few studies have developed multi-player models for the TAS problem in which terminal operators and drayage firms are non-cooperating stakeholders with conflicting objectives. Chen et al. (2013c) proposed a multi-player dynamic (i.e., same day appointments) TAS in which drayage firms make appointments based on actual waiting time. The terminal would deny the requested appointments if accepting them would result in a long queue; in this scenario, the drayage firm would need to request the appointment at some other time. They found that the proposed multi-player dynamic TAS can overcome the limitation of requiring arrival information in advance as an input to the single-player model. Phan and Kim (2015, 2016) proposed a multi-player decision-making model in which drayage firms and terminal operators make decisions independently. They modeled each drayage firm's preferences and constraints which were not considered in their previous work with the single-player model. They modeled the interplay between the drayage firms and terminal operator via an iterative process. In the iterative process, first each drayage firm requests appointments at its preferred times. Then the terminal operator predicts the queueing time after receiving all the requested appointments. Lastly, the terminal operator provides the assigned appointments to drayage firms. The drayage firms may keep or change their assigned appointments. This process is repeated until no additional changes are made by the drayage firms and terminal operator. The authors stated that "although the solution of multi-player model is worse than the single-player model, but the solution of multi-player model would reflect the real-life situation better than single-player model." Zhang et al. (2017) proposed a bi-level truck congestion pricing model to optimize toll rates at the terminal. The upper-level problem sought to minimize the average truck waiting time while the lower-level problem sought to minimize the drayage cost via a utility function. They designed a memetic heuristic to solve the bi-level problem. They found that the developed toll pricing model alleviated truck congestion and improved terminal efficiency.

Table 3.1 Summary of the TAS literature review

Author(s) (year)	Study country	TAS design									Solution method					TAS type				Key findings relevant to the current study
									Queuing system		Simulation model			Opt*						
		Q*	Y*	E*	A*	D*	V*	C/I*	S*	NS*	AG*	DE*	Qn*	Ex*	Hu*	SP*	MP*	TO*	RP*	
Morais and Lord (2006)	US*, CA*	✓		✓	✓				✓				✓					✓		Extended hours and TAS can reduce truck idling time
Huynh and Walton (2008)	US*	✓							✓			✓		✓				✓		Truckers and terminal would benefit from TAS
Huynh (2009)	US*	✓			✓				✓			✓						✓		TAS reduces truck turn time by 44%.
Guan and Liu (2009a;2009b)	US*	✓							✓					✓				✓		TAS can reduce the gate congestion.
Zhao and Goodchild (2010)	H*	✓	✓												✓			✓		Small information about arrivals reduces rehandles.

Table 3.1 Summary of the TAS literature review (Continued)

Chen and Yang (2010)	CN*	✓	✓	✓	✓	✓				✓		✓	Time-window optimization levels out truck arrivals.
Chen et al. (2011)	H*	✓	✓				✓	✓	✓		✓		PSFFA improves computational accuracy.
Chen et al. (2013a)	CN*	✓	✓				✓			✓		✓	TAS levels out arrivals and reduce the gate congestion.
Chen et al. (2013b)	US*	✓		✓	✓		✓			✓		✓	A small shift of arrivals can drastically reduce emissions.
Chen et al. (2013c)												✓	Dynamic TAS can increase the system flexibility.
Zhang et al (2013)	CN*	✓	✓				✓			✓		✓	TAS can decrease the truck turn time efficiently.

Table 3.1 Summary of the TAS literature review (Continued)

Zehendner and Feillet (2014)	FR*	✓						✓	✓		✓			✓	TAS can shift truck arrivals to off-peak hours.
Schulte et al. (2015)	CL*	✓		✓	✓	✓	✓		✓			✓			Collaborative TAS reduces emissions but might increase congestion.
Phan and Kim (2015)	H*	✓	✓			✓		✓		✓		✓	✓		Single-player and multi-player TASs differ by 3.03%.
Phan and Kim (2016)	H*	✓	✓		✓	✓		✓		✓			✓		The number of iterations is 9.2 on average.
Shiri and Huynh (2016)	H*	✓			✓		✓			✓	✓	✓			TAS can reduce drayage firms' operation time.

Table 3.1 Summary of the TAS literature review (Continued)

Chen and Jiang (2016)	CN*	✓	✓				✓		✓			✓		✓		TAS is effective way of managing truck arrivals.
Ramírez-Nafarrate et al. (2016)	CL*	✓	✓							✓				✓		TAS may reduce rehandles and truck waiting times.
Huynh et al. (2016)															✓	NA
Schulte et al. (2017)	CL*	✓	✓	✓	✓	✓		✓			✓		✓			TAS provides coordinated truck schedules.
Islam (2017)	NZ*			✓		✓		✓		✓			✓			TAS increases efficiency and reduces emissions.
Zhang and Zhang (2017)															✓	NA

Table 3.1 Summary of the TAS literature review (Continued)

Azab et al. (2017)	H*	✓	✓	✓	✓	✓	✓		✓				✓	TAS reduces the wait time and levels out the workload
Zhang et al. (2017)	CN*	✓			✓		✓				✓		✓	Pricing can decrease truck's queuing time effectively.
Torkjazi et al. (2018)	US*	✓	✓	✓	✓			✓		✓		✓		TAS reduces the drayage operation cost by 11.5%.
Jovanivic (2018)	US*	✓			✓					✓		✓		TAS reduces waiting times and benefits truck driver.
Lang et al. (2018)	DE*	✓			✓				✓			✓		The success of TAS depends on the pattern of drayage firms.

Table 3.1 Summary of the TAS literature review (Continued)

Li et al. (2018)	CN*	✓	✓							✓					✓	The proposed TAS reduces turn time about 76%.
Yang et al. (2018)	H*	✓	✓	✓	✓	✓					✓				✓	TAS improves yard efficiency.
Yi et al. (2019)	KR*	✓	✓		✓		✓		✓			✓		✓		The proposed TAS can reduce drayage cost by 15%
Current study	US*	✓		✓	✓			✓			✓		✓	✓		NA

*US: United States; *CA: Canada; *CN: China; *CL: Chile; *FR: France; *DE: Germany; *NZ: New Zealand; *KR: South Korea; *H: Hypothetical; *Q: Quotas considerations; *Y: Yard considerations; *E: Environmental considerations; *A: Assessment of existing TAS methods; *D: Drayage firms considerations; V*: Vessle arrival consideration; *C/I: Collaborative/Iterative TAS; *S: Stationary; *NS: Non-Stationary; *AG: Agent-Based; *DE: Discrete-Event; *Qn: Quantitative; *Opt: Optimization; *Ex: Exact; *Hu: Heuristic; *SP: Single-player; *MP: Multi-player; *TO: Terminal only; *RP: Review paper.

This study contributes to the body of work on developing multi-player models for the TAS problem. The specific contributions are: 1) modeling the TAS problem as a multi-player game and formulating it as a bi-level program with the terminal functions as the leader and drayage firms function as followers, and 2) converting the bi-level program to a single-level program by replacing the lower-level problem with its equivalent KKT conditions.

3.2 PROBLEM DESCRIPTION AND FORMULATION

The TAS problem is considered in the context of a marine container terminal servicing multiple ocean carriers and drayage firms. The terminal requires every truck to make an appointment by indicating its desired arrival time(s); the specific arrival times are considered by the TAS to be as accommodating as possible for each truck. Many terminals in the U.S. currently require trucks to make appointments such as those at the Port of Los Angeles and Long Beach. Every appointment request is associated with a container number and a truck ID. The appointment request needs to be submitted by 5:00 PM of the previous day (policy of Port of Vancouver). Similarly, the TAS requires the terminal operator to submit the quotas by 5 PM. Once the TAS received the appointment requests and quotas, it will calculate the time-gap between consecutive appointment requests for those trucks with more than one appointment. The time-gap is taken as the minimum time the truck needs to complete the consecutive jobs; that is, it is assumed that each truck will seek to complete as many jobs as possible in a given day. Furthermore, it is assumed that the TAS will have to inform drayage firms by 7:00 PM regarding their appointment requests (i.e., accepted or given another time-window). Note that although the truck specifies a desired arrival time when making the appointment, the TAS will provide the trucks with an appointment window and not a specific time to give trucks some flexibility in navigating traffic congestion, accidents and other potential delays. Figure 3.1 illustrates the described process.

It is assumed in the proposed TAS models that all trucks are required to make an appointment through the TAS which is in practice by Eagle Marine Services, eModal. It is assumed that the terminal gates have enough capacity to process all truck during their appointment time-window, so, no appointment will be transferred to next time-window. It is assumed that no-show appointments are required to request a new appointment for next day. It is also assumed that drayage firms do not share any information regarding their appointment requests.

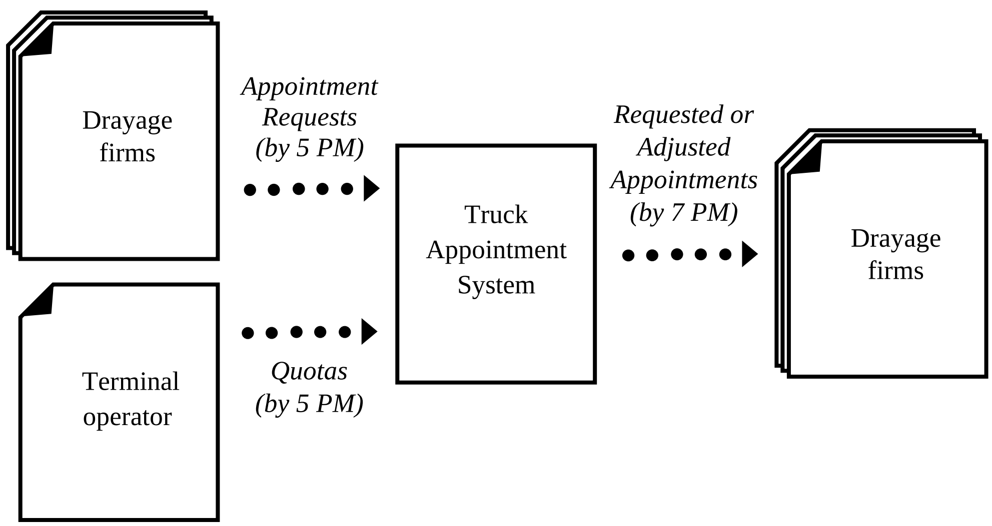


Figure 3.1 Appointment reservation process

3.2.1 FORMULATION OF THE MULTI-PLAYER TAS PROBLEM

To solve the described TAS problem, a bi-level optimization model is formulated. At the upper level, the leader (i.e., terminal) seeks to minimize the total gate waiting cost considering when each truck will arrive at the terminal. At the lower level, each drayage firm determines the best appointment window for each truck based on drayage cost and gate waiting cost. The optimal solution to this problem corresponds to the Stackelberg equilibrium. That is, the optimal solution yields the appointment times for each truck to achieve the lowest possible gate waiting cost and that the drayage companies cannot reduce their costs any further by unilaterally changing their appointment times.

In order to obtain the optimal equilibrium solution, the bi-level model needs to be converted into one single-level problem. This can be achieved by replacing the individual lower-level models by its equivalent KKT conditions. Such approach has been applied in the works of Asgari et al. (2013), Asadabadi and Miller-Hooks (2018) and Gao and You (2018). It should be noted that the lower-level (drayage) models, based on the work of Torkjazi et al. (2018), have binary variables, and therefore, the KKT conditions are not satisfied. In the following, three lemmas are introduced to enable the derivation of KKT conditions and to make the model tractable:

- Lemma 1: account for arrival penalties at time-windows in the lower-level problem, while keeping all the lower-level variables continuous to satisfy KKT conditions.
- Lemma 2: assign a discrete time-window to every continuous truck arrival time.
- Lemma 3: apply nonlinear penalty to number of arrivals in each time-window while keeping the problem linear.

3.2.1.1 LOWER-LEVEL DRAYAGE FIRM PROBLEM

The lower-level problem is formulated from the perspective of individual drayage firms. Each individual drayage firm determines truck arrival times with the goal of minimizing the drayage cost given gate waiting cost for each time-window. The lower-level problem along with its notations, parameters, decision variables and dual variables are presented in Table 3.2.

Table 3.2 Notation for the lower-level drayage firm problem

Sets		*Subsets*		*Indices*	
D	Drayage firms	K_d	Trucks from drayage firm d	t	Time-window
K	Trucks	A_i	Number of appointments for trucks with i appointments	j	Extra arrival
A	Appointment numbers			k	Truck
T	Time-windows			a	Appointment
J	Extra arrivals			d	Drayage firm

Parameters	
p_{ka}	Preferred arrival time of a^{th} appointment of truck k
C^{gap}_{large}	Penalty value for actual time-gap larger than the preferred time-gap
C^{gap}_{small}	Penalty value for actual time-gap smaller than the preferred time-gap
C_{pos}	Penalty value applied to truck that arrives earlier than scheduled
C_{neg}	Penalty value applied to truck that arrives later than scheduled
M_1	Auxiliary gate waiting cost effect (unit cost per gate waiting cost)
M_2	Penalty value for total appointment deviation from time-windows' mid-value (unit cost per gate waiting cost)
M_3	Gate waiting cost effect (gate waiting cost per appointment deviation from time-windows' mid-value)
n_t	Mid-point of time-window t
V_t	Gate waiting cost at time-window t

Decision variables	
X_{ka}	Arrival time of a^{th} appointment of truck k
N_{ka}	Difference between time-gap between a^{th} and $(a+1)^{th}$ appointments of truck k
Q_{ka}	Difference between time-gap between $(a+1)^{th}$ and a^{th} appointments of truck k
S_{ka}	Difference between actual and preferred arrival time-window
Z_{ka}	Difference between preferred and actual arrival time-window
g_{ka}	Surplus gate waiting cost variable
d^+_{kat}	Positive appointment deviation from time-windows' mid-value
d^-_{kat}	Negative appointment deviation from time-windows' mid-value

Dual variables
$\propto_{ka}, \beta_{ka}, \gamma_{ka}, \delta_{ka}, \omega_{ka}, \rho_{ka}, \eta_{ka}, \theta_{ka}, \lambda_{ka}, \mu_{ka}, \nu_{ka}, \pi_{kat}, \varphi_{kat}, \sigma_{kat}, \psi_{kat}, \varepsilon_{ka}, \chi$

$$\mathbf{Min}\ \ C_{large}^{gap}\sum_{k\in K_d}\sum_{a\in A_i}N_{ka}+C_{small}^{gap}\sum_{k\in K_d}\sum_{a\in A_i}Q_{ka}+ \tag{3.1}$$

$$C_{pos}\sum_{k\in K_d}\sum_{a\in A_i}S_{ka}+C_{neg}\sum_{k\in K_d}\sum_{a\in A_i}Z_{ka}+M_1\sum_{k\in K_d}\sum_{a\in A_i}g_{ka}+$$

$$M_2\sum_{k\in K_d}\sum_{a\in A_i}\sum_{t\in T}(d_{kat}^{+}\cdot d_{kat}^{-})$$

Eq. (3.1) is the objective function which seeks to minimize the total cost for one drayage firm. The first term of the equation is the cost of increasing the gap, from its preferred value, between two consecutive appointments. The second term is the cost of making the gap smaller than the preferred gap for those trucks with more than one appointment. The third term is the cost of shifting an appointment to a later time-window and the fourth term is the cost of shifting an appointment to an earlier time-window. The fifth term represent the gate waiting cost. The sixth term is a penalty cost, required to ensure that one of the d_{kat}^{+} and d_{kat}^{-} is zero for each k, a, and t. Each d_{kat}^{+}, d_{kat}^{-} pair would determine the arrival time distance from the center of a specific time-window (explained later in lemma1). The objective function (Eq. (3.1)) is subject to constraints (3.2) to (3.16):

		Duals	
$-N_{ka}\le 0$	$\forall k\in K_d, \forall a\in A_i\setminus\{i\}: i>1$	$:\propto_{ka}$	(3.2)
$X_{k(a+1)}-X_{ka}-N_{ka}+p_{ka}-p_{k(a+1)}\le 0$	$\forall k\in K_d, \forall a\in A_i\setminus\{i\}: i>1$	$:\beta_{ka}$	(3.3)
$-Q_{ka}\le 0$	$\forall k\in K_d, \forall a\in A_i\setminus\{i\}: i>1$	$:\gamma_{ka}$	(3.4)
$-Q_{ka}+p_{k(a+1)}-p_{ka}+X_{ka}-X_{k(a+1)}\le 0$	$\forall k\in K_d, \forall a\in A_i\setminus\{i\}: i>1$	$:\delta_{ka}$	(3.5)

$$-S_{ka} \leq 0 \qquad \forall k \in K_d, \forall a \in A_i \qquad : \omega_{ka} \qquad (3.6)$$

$$X_{ka} - p_{ka} - S_{ka} \leq 0 \qquad \forall k \in K_d, \forall a \in A_i \qquad : \rho_{ka} \qquad (3.7)$$

$$-Z_{ka} \leq 0 \qquad \forall k \in K_d, \forall a \in A_i \qquad : \eta_{ka} \qquad (3.8)$$

$$-Z_{ka} - X_{ka} + p_{ka} \leq 0 \qquad \forall k \in K_d, \forall a \in A_i \qquad : \theta_{ka} \qquad (3.9)$$

Constraints (3.2), (3.4), (3.6) and (3.8) are non-negativity constraints. Given a desired gap between two appointments for a truck, constraint (3.3) calculates the increase over the desired gap with the assigned time-windows, whereas constraint (3.5) calculates the reduction of the desired gap with the assigned time-windows. Similarly, given a desired time-window for an appointment, constraints (3.7) and (3.9) calculate the difference between the assigned and desired arrival time. To prevent trucks from arriving in a manner that create congestion, a penalty mechanism is employed as explained in Lemma 1.

<u>Lemma 1:</u> *Penalizing a truck arrival within a time-window using continuous variables.*

<u>*Duals*</u>

$$X_{ka} - n_t = d^{+}_{kat} - d^{-}_{kat} \qquad \forall k \in K_d, \forall a \in A_i, \forall t \in T \qquad : \pi_{kat} \qquad (3.10)$$

$$V_t - M_3(d^{+}_{kat} + d^{-}_{kat}) \leq g_{ka} \qquad \forall k \in K_d, \forall a \in A_i, \forall t \in T \qquad : \varphi_{kat} \qquad (3.11)$$

Eq. (3.10) employs two variables d^{+}_{kat} and d^{-}_{kat} to find the absolute value of the time difference between the actual arrival time, X_{ka}, and the mid-point of each time-window, n_t; A truck arrival is either greater than, less than or at the time-window's mid-point. The multiplication of these two variables is associated with the big M in the objective function (sixth term of the objective function) to ensure at least one of them will be zero. This condition is added to the objective function to keep the constraint set linear and to allow for determination of KKT equality conditions. The estimated d_{kat}^{+} and d_{kat}^{-} are used in constraint (3.11) to penalize each arrival according to the penalty of its specific time-window, V_t. According to this constraint, if a truck is arriving at the mid-point of a time-window (e.g., 2:30 PM for the time-window from 2 PM to 3 PM), it will incur the maximum penalty. This penalty is reduced linearly the further away the truck's arrival time is to the mid-point of the time window. The surplus gate waiting cost, g_{ka}, will be minimized in the objective function which encourages the drayage companies to avoid busy time-windows and associated waiting times.

The time-window duration constraint is expressed as follows:

		Dual	
$X_{ka} - q \leq 0$	$\forall k \in K_d, \forall a \in A_i$	$: \lambda_{ka}$	(3.12)

Constraint (3.12) limits the truck arrival time to a pre-specified time, q. Lastly, constraints (3.13) to (3.16) define the domain of decision variables.

		Duals	
$-X_{ka} \leq 0$	$\forall k \in K_d, \forall a \in A_i$	$: \nu_{ka}$	(3.13)
$-d_{kat}^{+} \leq 0$	$\forall k \in K_d, \forall a \in A_i, \forall t \in T$	$: \sigma_{kat}$	(3.14)
$-d_{kat}^{-} \leq 0$	$\forall k \in K_d, \forall a \in A_i, \forall t \in T$	$: \psi_{kat}$	(3.15)
$-g_{ka} \leq 0$	$\forall k \in K_d, \forall a \in A_i$	$: \varepsilon_{ka}$	(3.16)

3.2.1.2 UPPER-LEVEL TERMINAL OPERATOR PROBLEM

In the upper-level problem, the terminal operator decides on penalties for each arrival time period with the goal of spreading out the truck arrivals while anticipating the response of the drayage companies. The term "penalties" in this model are not actual fees or costs that will be incurred by the trucks if they do not arrive at the assigned time-windows; rather, they represent waiting cost that is used as a measure to increase or decrease the utility (i.e., attractiveness) of a time-window. The objective function minimizes the total gate waiting cost for all time-windows; in other words, it seeks to minimize the total waiting cost for all trucks. The upper-level problem along with its notations, parameters and decision variables are presented in Table 3.3.

Table 3.3 Notation for the upper-level terminal operator problem

Sets		*Subsets*		*Indices*	
D	Drayage firms	K_d	Trucks from drayage firm d	t	Time-window
K	Trucks	A_i	Number of appointments for trucks with i appointments	j	Extra arrival
A	Appointment numbers			k	Truck
T	Time-windows			a	Appointment
J	Extra arrivals			d	Drayage firm

Parameters	
C_t	Quota at time-window t
X_{ka}	Arrival time of a^{th} appointment of truck k

Decision variables	
B_{kat}	1 if truck k has its a^{th} appointment at time-window t, 0 otherwise
D_{tj}	1 if time-window t has more than j arrivals, 0 otherwise
Y_t	Number of arrivals at time-window t
U_{tj}	Penalty value for j^{th} arrivals at time-window t
V_t	Gate waiting cost at time-window t

The objective function of the upper-level problem is presented as follows:

$$\mathbf{Min} \sum_{t \in T} V_t \tag{3.17}$$

Eq. (3.17) minimizes the total gate waiting costs over all time-windows. It should be noted that in the bi-level structure of the model, the gate waiting cost is a decision variable in the upper-level problem, but it is an input parameter in the lower-level problem. Since the gate waiting cost variable will be determined according to the number of arrivals during each time-window, it is necessary to relate the continuous arrival variable, X_{ka}, to the binary variable, B_{kat}. This is achieved by Lemma 2.

Lemma 2: *Assigning a continuous truck arrival time to a discrete time-window.*

$$\sum_{t\in T}(t-1)\cdot B_{kat} \le X_{ka} \le \sum_{t\in T} t\cdot B_{kat} \qquad \forall k \in K_d, \forall a \in A_i \tag{3.18}$$

$$\sum_{t\in T} B_{kat} = 1 \qquad \forall k \in K_d, \forall a \in A_i, \forall t \in T \tag{3.19}$$

Constraint (3.18) finds the time-window that an arrival time falls into. Constraint (3.19) ensures that each appointment is assigned only to one time-window. The order of the appointments for a truck and the quota constraint are satisfied using constraints (3.20) and (3.21). Eq. (3.22) calculates the number of arrivals in each time-window.

$$\sum_{t\in T} t\cdot B_{kat} \le \sum_{t\in T} t\cdot B_{k(a+1)t} \qquad \forall k \in K_d, \forall a \in A_i\backslash\{i\}: i > 1 \tag{3.20}$$

$$\sum_{k\in K_d}\sum_{a\in A_i} B_{kat} \le C_t \qquad \forall t \in T \tag{3.21}$$

To spread out truck arrivals throughout the day a nonlinear penalty structure is utilized while preserving the linearity of the model. This is achieved by Lemma 3.

Lemma 3: *Assigning nonlinear penalty for number of arrivals while keeping the problem linear*

$$Y_t = \sum_{k\in K_d}\sum_{a\in A_i} B_{kat} \qquad \forall t \in T \tag{3.22}$$

$$Y_{t,j} \le M\left(D_{t,j}\right) + j - 1 \qquad \forall t \in T, \forall j \in J \tag{3.23}$$

$$\sum_{j\in J} U_{tj}\cdot D_{tj} \le V_t \qquad \forall t \in T \tag{3.24}$$

To obtain nonlinear penalty for the number of arrivals in each time period in a linear formulation a set of penalties, U_{tj}, are calculated that represents the additional penalty of one more arrival over j at time t. U_{tj} is a linear function of j, which mean that $\sum_{j\in J} U_{tj}$ is a nonlinearly increasing function. Constraint (3.23) would force $D_{t,j}$ to one when the number of arrivals at time t is more than j to ensure that U_{tj} is active in constraint (3.24). Constraints (3.25) and (3.26) define the domain of the decision variables.

$$D_{t,j} \in \{0,1\} \qquad \forall t \in T, \forall j \in J \qquad (3.25)$$

$$B_{kat} \in \{0,1\} \qquad \forall k \in K_i, \forall a \in A_i, \forall t \in T \qquad (3.26)$$

3.2.1.3 SINGLE-LEVEL MULTI-PLAYER TAS PROBLEM

The bi-level, multi-player programming problem of terminal operator and drayage firms can be summarized as follows:

$$\begin{aligned}
&\textbf{Min } (3.17) \\
&\quad \text{subject to:} \\
&\quad (3.18) - (3.26) \\
&\quad \textbf{Min } (3.1) \\
&\qquad \text{subject to:} \\
&\qquad (3.2) - (3.16)
\end{aligned}$$

This mathematical formulation falls in the category of Optimization Problems constrained with other Optimization Problems (OPcOPs) (Gabriel et al., 2012). In order to solve this set of connected optimization problems, the OPcOP is transformed into a single-level model, which is later linearized and can be solved to optimality. To accomplish this, the lower-level models of drayage companies, (2)-(16), are replaced by their KKT equivalent conditions, resulting in a Mathematical Problem with Equilibrium

Constraints (MPEC) (Gabriel et al., 2012). Since the modified lower-level problems via Lemma 1 contain only continuous variables and linear constraints and the objective functions is convex, the KKT conditions for each drayage company are necessary and sufficient for optimality, and the bi-level model reduces to single-level model as follows:

$\textbf{Min}$ (3.17)

subject to:

$(3.18) - (3.26)$

KKT CONDITIONS FOR (3.2)-(3.16)

The KKT conditions for equations (3.2)-(3.16) including primal feasibility, stationarity, dual feasibility, and complementary slackness are as follows:

$$\frac{\partial L}{\partial X_{ka}} = \rho_{ka} - \theta_{ka} + \lambda_{ka} + \mu_{ka} - \nu_{ka} + \sum_{t \in T} \pi_{kat} = 0 \qquad \forall k \in K_d, a = 1: a \in A_i \;\&\; i = 1 \tag{3.27}$$

$$\frac{\partial L}{\partial X_{ka}} = -\beta_{ka} + \delta_{ka} + \rho_{ka} - \theta_{ka} + \lambda_{ka} + \mu_{ka} - \nu_{ka} + \sum_{t \in T} \pi_{kat} = 0 \qquad \forall k \in K_d, a = 1: a \in A_i \;\&\; i > 1 \tag{3.28}$$

$$\frac{\partial L}{\partial X_{ka}} = \beta_{k(a-1)} - \beta_{ka} - \delta_{k(a-1)} + \delta_{ka} + \rho_{ka} - \theta_{ka} + \lambda_{ka} - \mu_{k(a-1)} + \mu_{ka} - \nu_{ka} + \sum_{t \in T} \pi_{kat} = 0 \qquad \forall k \in K_d, 1 < a < i: a \in A_i \;\&\; i > 1 \tag{3.29}$$

$$\frac{\partial L}{\partial X_{ka}} = \beta_{k(a-1)} - \delta_{k(a-1)} + \rho_{ka} - \theta_{ka} + \lambda_{ka} - \mu_{k(a-1)} - \nu_{ka} + \sum_{t \in T} \pi_{kat} = 0 \qquad \forall k \in K_d, a = i: a \in A_i \;\&\; i > 1 \tag{3.30}$$

$$\frac{\partial L}{\partial N_{ka}} = -C_{large}^{gap} + \alpha_{ka} + \beta_{ka} = 0 \qquad \forall k \in K_d, \forall a \in A_i \backslash \{i\}: i > 1 \tag{3.31}$$

$$\frac{\partial L}{\partial Q_{ka}} = -C_{small}^{gap} + \gamma_{ka} + \delta_{ka} = 0 \qquad \forall k \in K_d, \forall a \in A_i \backslash \{i\}: i > 1 \qquad (3.32)$$

$$\frac{\partial L}{\partial S_{ka}} = -C_{pos} + \omega_{ka} + \rho_{ka} = 0 \qquad \forall k \in K_d, \forall a \in A_i \qquad (3.33)$$

$$\frac{\partial L}{\partial Z_{ka}} = -C_{neg} + \eta_{ka} + \theta_{ka} = 0 \qquad \forall k \in K_d, \forall a \in A_i \qquad (3.34)$$

$$\frac{\partial L}{\partial d_{kat}^{+}} = -M_2 \cdot d_{kat}^{-} + \pi_{kat} + M_3 \cdot \varphi_{kat} + \sigma_{kat} = 0 \qquad \forall k \in K_d, \forall a \in A_i, \forall t \in T \qquad (3.35)$$

$$\frac{\partial L}{\partial d_{kat}^{-}} = -M_2 \cdot d_{kat}^{+} - \pi_{kat} + M_3 \cdot \varphi_{kat} + \psi_{kat} = 0 \qquad \forall k \in K_d, \forall a \in A_i, \forall t \in T \qquad (3.36)$$

$$\frac{\partial L}{\partial g_{ka}} = -M_1 + \sum_{t \in T} \varphi_{kat} + \varepsilon_{ka} = 0 \qquad \forall k \in K_d, \forall a \in A_i \qquad (3.37)$$

$$0 \leq \alpha_{ka} \perp N_{ka} \geq 0 \qquad \forall k \in K_d, \forall a \in A_i \backslash \{i\}: i > 1 \qquad (3.38)$$

$$0 \leq \beta_{ka} \perp -X_{k(a+1)} + X_{ka} + N_{ka} - p_{ka} + p_{k(a+1)} \geq 0 \qquad \forall k \in K_d, \forall a \in A_i \backslash \{i\}: i > 1 \qquad (3.39)$$

$$0 \leq \gamma_{ka} \perp Q_{ka} \geq 0 \qquad \forall k \in K_d, \forall a \in A_i \backslash \{i\}: i > 1 \qquad (3.40)$$

$$0 \leq \delta_{ka} \perp Q_{ka} - p_{k(a+1)} + p_{ka} - X_{ka} + X_{k(a+1)} \geq 0 \qquad \forall k \in K_d, \forall a \in A_i \backslash \{i\}: i > 1 \qquad (3.41)$$

$$0 \leq \omega_{ka} \perp S_{ka} \geq 0 \qquad \forall k \in K_d, \forall a \in A_i \qquad (3.42)$$

$$0 \leq \rho_{ka} \perp -X_{ka} + p_{ka} + S_{ka} \geq 0 \qquad \forall k \in K_d, \forall a \in A_i \qquad (3.43)$$

$$0 \leq \eta_{ka} \perp Z_{ka} \geq 0 \qquad \forall k \in K_d, \forall a \in A_i \qquad (3.44)$$

$$0 \leq \theta_{ka} \perp Z_{ka} + X_{ka} - p_{ka} \geq 0 \qquad \forall k \in K_d, \forall a \in A_i \qquad (3.45)$$

$$0 \leq \lambda_{ka} \perp -X_{ka} + q \geq 0 \qquad \forall k \in K_d, \forall a \in A_i \qquad (3.46)$$

$$0 \leq \mu_{ka} \perp -X_{ka} + X_{k(a+1)} \geq 0 \qquad \forall k \in K_d, \forall a \in A_i \backslash \{i\}: i > 1 \qquad (3.47)$$

$$0 \leq \nu_{ka} \perp X_{ka} \geq 0 \qquad \forall k \in K_d, \forall a \in A_i \qquad (3.48)$$

$$0 \leq \varphi_{kat} \perp -V_t + M_3(d_{kat}^{+} + d_{kat}^{-}) + g_{ka} \geq 0 \qquad \forall k \in K_d, \forall a \in A_i, \forall t \in T \qquad (3.49)$$

$$0 \leq \sigma_{kat} \perp d_{kat}^{+} \geq 0 \qquad \forall k \in K_d, \forall a \in A_i, \forall t \in T \qquad (3.50)$$

$$0 \leq \psi_{kat} \perp d_{kat}^{-} \geq 0 \qquad \forall k \in K_d, \forall a \in A_i, \forall t \in T \qquad (3.51)$$

$$0 \leq \varepsilon_{ka} \perp g_{ka} \geq 0 \qquad \forall k \in K_d, \forall a \in A_i \qquad (3.52)$$

The function $\perp$ in constraint (3.43) is used to indicate that it is equivalent to the following:

$0 \leq \rho_{ka}$	Non-negativity of the dual variable
$0 \leq -X_{ka} + p_{ka} + S_{ka}$	A lower-level inequality
$\rho_{ka}(-X_{ka} + p_{ka} + S_{ka}) = 0$	Complementary slackness

A disjunctive constraint approach (Fortuny-Amat and McCarl, 1981) is adopted in creating equivalent linear constraint for complementary slackness constraints (3.38)-(3.52). The resulting TAS problem from using this linearization technique, is a Mixed

Integer programming (MIP) problem. As an example, the linearization of constraint (3.43) follows:

$$-X_{ka} + p_{ka} + S_{ka} \leq (1 - a)M_1 \tag{3.53}$$

$$\rho_{ka} \leq (a)M_2 \tag{3.54}$$

Where M_1 and M_2 are large values that place no restrictions on $(-X_{ka} + p_{ka} + S_{ka})$ and ρ_{ka} when a is 1 or 0, respectively.

3.1.2 Single-level Single-player TAS Problem

For purpose of comparison, an adapted version of the single-player TAS from the work of Torkjazi et al. (2018) is utilized. As mentioned, in the single-player version of the TAS, a single decision-maker solves the truck appointment problem with the objective of minimizing the sum of the lower-level and upper-level objective functions, subject to the combined set of constraints of the lower-level and upper-level problems. The single-player version of the TAS problem can be expressed as follows:

$$\textbf{Min} \quad C^{gap}_{large} \sum_{k \in K_d} \sum_{a \in A_i} N_{ka} + C^{gap}_{small} \sum_{k \in K_d} \sum_{a \in A_i} Q_{ka} + \tag{3.55}$$

$$C_{pos} \sum_{k \in K_d} \sum_{a \in A_i} S_{ka} + C_{neg} \sum_{k \in K_d} \sum_{a \in A_i} Z_{ka} + C_{arr} \sum_{t \in T} V_t$$

subject to:

$(3.2) - (3.9)$

$(3.12) - (3.13)$

$(3.18) - (3.26)$

Where, C_{arr} is the coefficient of the gate waiting cost component. To avoid having a nonlinear term in the objective function of the single-player model, the last term of the Eq. (3.1) is removed from the objective function of the single-player model and is replaced with equivalent constraints (3.56) and (3.57). These two constraints ensure at least one of two variables d^{+}_{kat} and d^{-}_{kat} is zero.

$$d^{+}_{kat} \leq n^{r}_{kat} \cdot M_4 \qquad \forall k \in K_d, \forall a \in A_i, \forall t \in T \tag{3.56}$$

$$d^{-}_{kat} \leq (1 - n^{r}_{kat}) \cdot M_4 \qquad \forall k \in K_d, \forall a \in A_i, \forall t \in T \tag{3.57}$$

Where, n^{r}_{kat} is an auxiliary binary variable and M_4 is a big number.

3.3 ILLUSTRATIVE EXAMPLE

Due to the complexity of the proposed single-player and multi-player models, their inner workings are illustrated via an illustrative example. This example is aimed to show: 1) how each model works, 2) why a multi-player TAS can model the decision-making behavior of terminal and trucking companies much more realistically, and 3) how moving from a single-player model to a multi-player model can produce different and more realistic results. Table 3.4 provides a summary of the parameters used for this illustrative example, which has three one-hour time-windows and two drayage firms with each having one truck. The first time-window is assumed to start at time zero. Truck 1 belongs to drayage firm 1, and it requested 2 appointments at time 0.99 and 1.5 which falls into time-windows 1 and 2, respectively. Truck 2 belongs to drayage firm 2, and it requested one appointment at time 1.5 which is associated with time-window 2.

The penalty for the n^{th} extra arrival at a time-window is assumed to be an increasing function as shown in Eq. (3.58). This structure ensures a penalty increase for each additional arrival beyond the quota in a time-window. As shown in Eq. (3.58), the gate waiting cost for arriving at a certain time-window is a function of number of trucks that will arrive at that time-window.

$$V_t = \sum_{Y_t} 10(Y_t) \tag{3.58}$$

The solution of the single-level, single-player TAS model is shown in Figure 3.2, where the truck preferred arrival times (initial requests/state 1) and the resulting single-player solution (optimal solution/state 2) are indicated by the arrows. As shown, there is one preferred arrival in time-window 1 and two preferred arrivals in time-window 2. The associated gate waiting costs for these arrivals using Eq. (3.58) are $(10) \cdot (1) = 10, (10) \cdot (1) + (10) \cdot (2) = 30,$ and $(10) \cdot (0) = 0$. The gate waiting costs are shown in dash circles. Thus, the objective function value of the single-player TAS model with preferred arrival times (state 1) via Eq. (3.55) is 10 + 30 + 0 = 40. In state 2, the single-player TAS shifts the arrival time of truck 2 to the earliest time of the next time-window (i.e., time 2.01 of time-window 3) to reduce the total gate waiting cost. This solution yields an objective function value of 35 (optimal solution), which is calculated as follows via Eq. (3.55): $C_{pos} \sum_{k \in K_d} \sum_{a \in A_i} S_{ka} + C_{arr} \sum_{t \in T} V_t = 35$.

In contrast to the single-player TAS solution, the solution of the multi-player TAS is shown in Figure 3.3. Based on number of preferred arrivals in each time-window, the terminal sets the gate waiting costs for those time-windows to minimize total gate waiting cost for all trucks. The gate waiting cost, from the drayage companies' perspective, are shown in dash circles. Note that the gate waiting cost at the mid-point of time-window 1 is 10, and it decreases linearly to 5 the further away the arrival time is to the mid-point (Eq. (3.11)). The same penalty pattern can be observed for time-window 2, except that the penalty cost at the mid-point is 30 instead of 10 because there are two arrivals in time-window 2 compared to one arrival in time-window 1 (Eq. (3.58)). The first appointment of truck 1 is at the end of time-window 1, so it has a gate waiting cost of 5, while its second appointment is at the mid-point of time-window 2 which has a gate waiting cost of 30. The gate waiting cost of truck 2's appointment is 30 since its appointment is at the mid-point of time-window 2. The best possible move is for truck 1 to shift its second appointment to the earliest time in next time-window (2.01) which would yield the lowest total gate waiting cost for its appointments. Note that this solution is different from that of the single-player TAS. After this move from the follower (drayage firm 1), the leader (terminal) updates the gate waiting costs via Eq. (3.58) to be 10 and 5 at the mid-point and end of every time-window (State 2). With the updated gate waiting cost, the objective function value of drayage firm 1 is calculated via Eq. (3.1) as $\mathrm{C}_{\mathrm{large}}^{\mathrm{gap}} \sum_{\mathrm{k} \in \mathrm{K}_{\mathrm{d}}} \sum_{\mathrm{a} \in \mathrm{A}_{\mathrm{i}}} \mathrm{N}_{\mathrm{ka}} + \mathrm{C}_{\mathrm{pos}} \sum_{\mathrm{k} \in \mathrm{K}_{\mathrm{d}}} \sum_{\mathrm{a} \in \mathrm{A}_{\mathrm{i}}} \mathrm{S}_{\mathrm{ka}} + \mathrm{M}_1 \sum_{\mathrm{k} \in \mathrm{K}_{\mathrm{d}}} \sum_{\mathrm{a} \in \mathrm{A}_{\mathrm{i}}} \mathrm{g}_{\mathrm{ka}} = 5 + 5 + 5 + 5 = 20.$ Similarly, the objective function value of drayage firm 2 is calculated as $\mathrm{M}_1 \sum_{\mathrm{k} \in \mathrm{K}_{\mathrm{d}}} \sum_{\mathrm{a} \in \mathrm{A}_{\mathrm{i}}} \mathrm{g}_{\mathrm{ka}} = 5 + 5 = 10$. At this point, no other changes by the drayage firms can reduce their objective function values. Therefore, the equilibrium optimal solution of the multi-player TAS is achieved.

Table 3.4 Parameters of illustrative example.

(1)	Total requested appointments	3
(2)	Set of drayage firms (D)	{1,2}
(3)	Set of trucks (K)	{1,2}
(4)	Set of trucks from drayage firm 1 (K_1)	{1}
(5)	Set of trucks from drayage firm 2 (K_2)	{2}
(6)	Set of time-windows (T)	{1,…,3}
(7)	M1	1
(8)	M2	100
(9)	M3	10
(10)	Quotas per time-window (C_t)	(2,2,2,2,2)
(11)	C^{gap}_{large}	10
(12)	C^{gap}_{small}	30
(13)	C_{pos}	10
(14)	C_{neg}	30

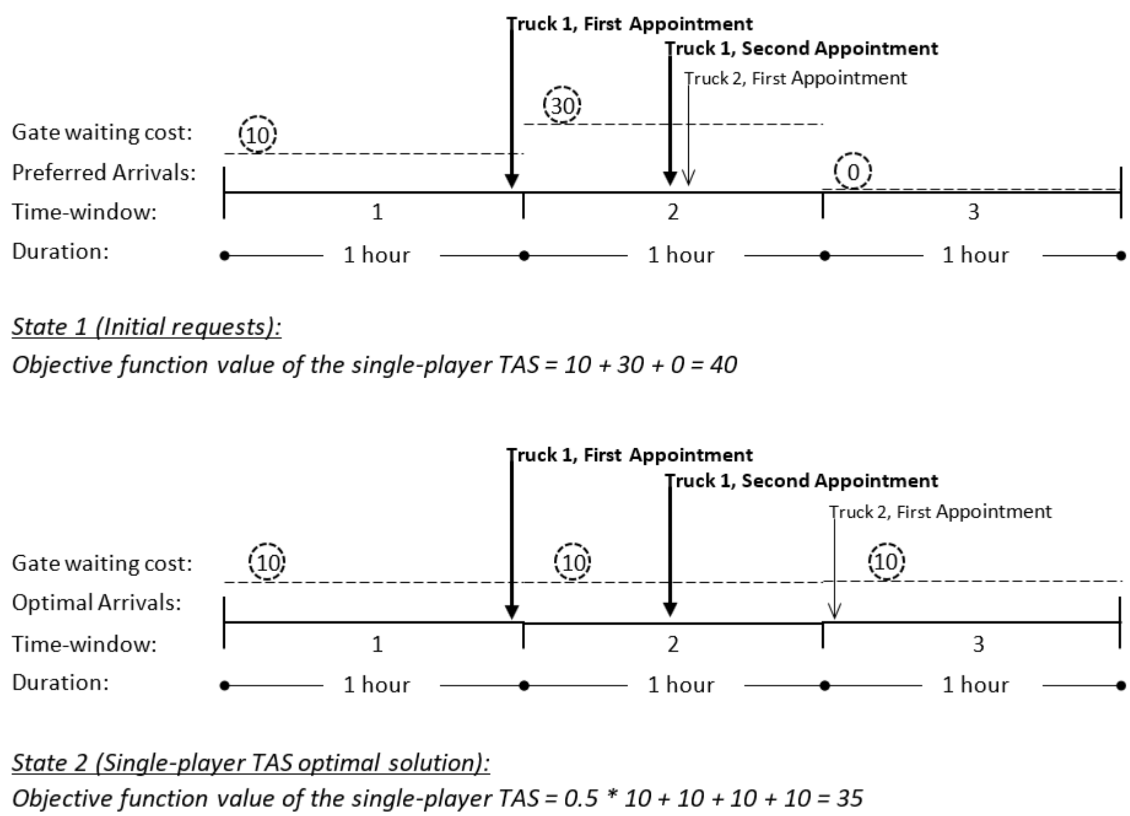


Figure 3.2 Solution of illustrative example using single-player TAS

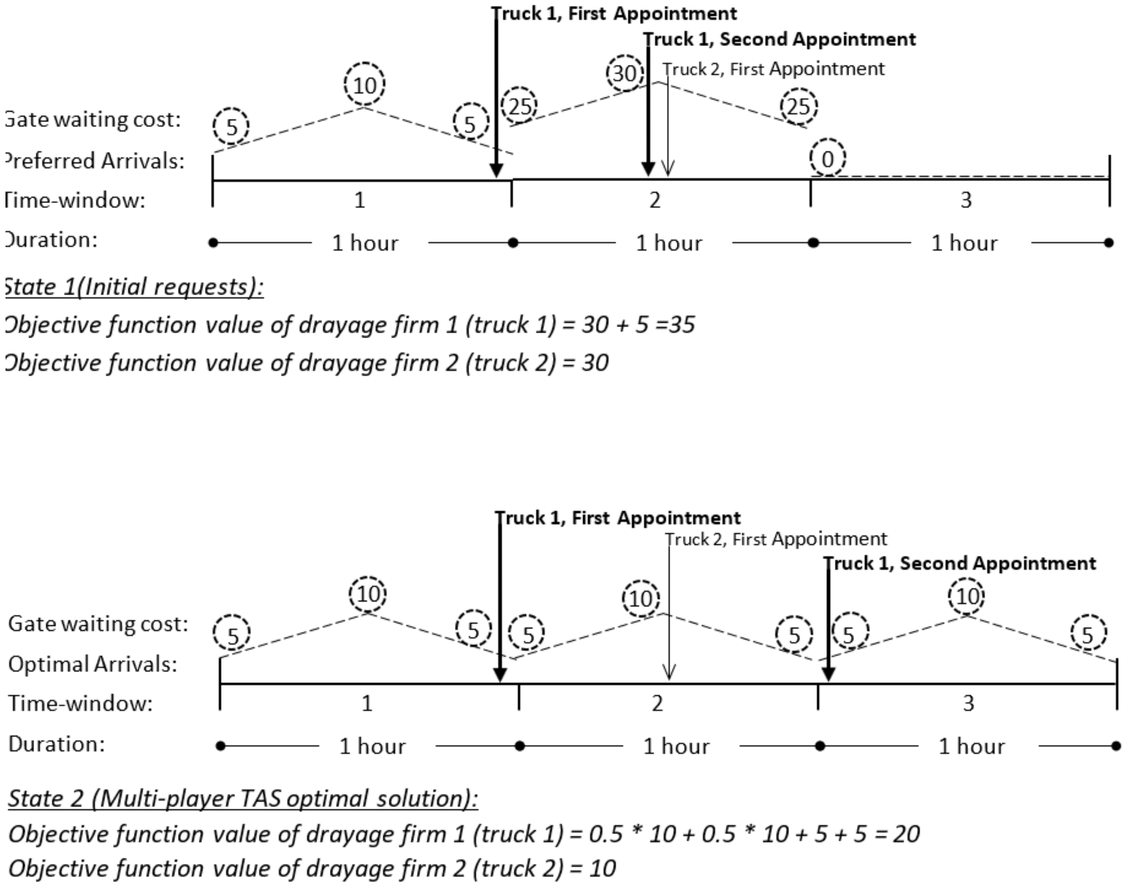


Figure 3.3 Solution of illustrative example using multi-player TAS

The above illustrative example shows that when different players are treated as individual decision makers, they do not necessarily make the same decisions as that made by a single decision maker. Differences between these two types of models (single-player vs. multi-player) are further explored through larger scenarios as described in the next section.

3.4 EXPERIMENTAL DESIGN

A series of experiments were designed to investigate the differences between the multi-player TAS model and single-player TAS model solutions in terms of drayage cost and waiting cost. Table 3.5 provides a summary of the parameters used for experiments 1 to 68. Experiments 1 to 50 aim to understand the effect of problem size (number of appointments varied between 4 to 22). These same set of experiments were used to examine how the solution of the multi-player TAS model differ from the single-player TAS model with different C_{arr} parameter values (between 0.01 and 1.0), each representing a different weight for the gate waiting cost component (Eq. (3.55)). The purpose of varying C_{arr} is to investigate the effect of gate waiting cost relative to drayage cost of the single-player model solutions (see Eq. 55). It should be noted that the multi-player TAS model has only one term in its objective function with a coefficient of one (see Eq. 17); its optimal solution is the same regardless of the coefficient value. Experiments 51 to 68 were designed to examine the effect of the average number of appointments per truck tour on the drayage cost (varied between 1 to 3). These experiments were conducted for three problem sizes: 6 appointments (experiments 51 to 56), 12 appointments (experiments 57 to 62), and 18 appointments (experiments 63 to 68). Preliminary experiments were performed to determine the largest problem size that could be solved within a day (24 hours) using the CPLEX (version 12.8) solver, and this was 22. For this reason, in designing the experiments, the largest problem size was kept at 22 appointments. All experiments were conducted on a desktop computer with Intel Core i7 3.4 GHz CPU and 16 GB of RAM.

Since the objective function of the single-level multi-player model is not comparable to the objective function of the single-player model, two separate terms are post-calculated from the results of both models for comparison purposes. The first term is equal to the summation of first four terms of Eq. (3.55) which is drayage cost. The other term is equal to the last term of Eq. (3.55) which is gate waiting cost.

The following discusses the model parameters used in the experiments conducted to examine the differences and advantages of utilizing the proposed multi-player model as compared to the single-player model. The developed single-level multi-player TAS model contains seven different penalty parameters. Their values are set based on a set preliminary experiments to ensure the model produces feasible and optimal results and that the results correspond to expectation. Values of the four drayage cost penalty parameters, C_{large}^{gap}, C_{small}^{gap}, C_{pos}, and C_{neg} are set to 10, 30, 10, and 30, respectively; these values are similar to the values used in the work of Torkjazi et al. (2018). The parameter M_1 ensures the penalty for arriving is applied at the actual arrival time-window (not the previous and next time-windows). M_2 should be large enough to ensure the multiplication of auxiliary variables ($d_{kat}^{+} \cdot d_{kat}^{-}$) is zero for each d_{kat}^{+} , d_{kat}^{-} combination, meaning that at least one of these two values will be zero. This parameter should have the highest value among multipliers of lower-level objective function to ensure feasibility. M_1, M_2 and M_3 are assigned values of 1, 100 and 30, respectively.

The experiments were generated using a combination of actual terminal data and data used in published studies. Each appointment time-window was considered to be 2 hours which is the case at the Port of Los Angeles and Long Beach. The terminal was assumed to operate 10 hours per day from 8:00 AM to 6:00 PM (Torkjazi et al., 2018). Preferred truck arrival times were generated randomly with the study area being equivalent to the size of the Port of Los Angeles and Long Beach region: a 3-hour by 3-hour (travel time), square area. The average turn time at the Port of LA/LB in November and December of 2016 were 45.2 and 42.9 minutes, respectively (PierPass). The average of these two values (44.05 min) was used for the experiments in this paper. It was assumed that the average gate queuing time at terminal is 10 minutes (Shiri and Huynh, 2016; Torkjazi et al., 2018). Quotas were assumed to be set at 2 times of the number of requested appointments (Shiri and Huynh, 2016; Torkjazi et al., 2018).

Table 3.5 Parameters of the TAS models

Experiment number	Size	Avg. number of appointments per truck	C_{arr}	Experiment number	Size	Avg. number of appointments per truck	C_{arr}
1	4	2	NA	35	16	2	1
2	4	2	0.01	36	18	2	NA
3	4	2	0.1	37	18	2	0.01
4	4	2	0.5	38	18	2	0.1
5	4	2	1	39	18	2	0.5
6	6	2	NA	40	18	2	1
7	6	2	0.01	41	20	2	NA
8	6	2	0.1	42	20	2	0.01
9	6	2	0.5	43	20	2	0.1
10	6	2	1	44	20	2	0.5
11	8	2	NA	45	20	2	1
12	8	2	0.01	46	22	2	NA
13	8	2	0.1	47	22	2	0.01
14	8	2	0.5	48	22	2	0.1
15	8	2	1	49	22	2	0.5
16	10	2	NA	50	22	2	1
17	10	2	0.01	51	6	1	NA
18	10	2	0.1	52	6	1	1
19	10	2	0.5	53	6	1	NA
20	10	2	1	54	6	1	1
21	12	2	NA	55	6	1	NA
22	12	2	0.01	56	6	1	1
23	12	2	0.1	57	12	2	NA
24	12	2	0.5	58	12	2	1
25	12	2	1	59	12	2	NA
26	14	2	NA	60	12	2	1
27	14	2	0.01	61	12	2	NA
28	14	2	0.1	62	12	2	1
29	14	2	0.5	63	18	3	NA
30	14	2	1	64	18	3	1
31	16	2	NA	65	18	3	NA
32	16	2	0.01	66	18	3	1
33	16	2	0.1	67	18	3	NA
34	16	2	0.5	68	18	3	1

3.5 EXPERIMENT RESULTS AND DISCUSSION

Table 3.6 shows the objective function value and run time of experiments 1 to 50. Each row of the table shows the results of 5 experiments; one for multi-player and four for single-player model with different C_{arr} values. It can be seen that the objective function value of multi-player and single-player models increase with problem size (columns 2, 4, 5, 6, and 7). The run time of the multi-player and single-player models also increase with problem size. For the multi-player model, its run time grows exponentially (column 3). Figure 3.4a shows graphically the impact of problem size on drayage cost. Note that the results of the single-player model are shown for different weights of the gate waiting cost component in Equation 3.55 (i.e., C_{arr}). When the weight is nearly zero, it represents the scenario where the single-player model assigns appointments primarily to lower drayage cost, whereas when the weight is 1, it represents the scenario where the single-player model assigns appointments strictly to lower gate waiting cost. As expected, the total drayage cost generally increases as the problem size increases. In some cases, the drayage cost remains the same or decreases slightly. The reason is due to random appointment request times.

Table 3.6 Results of experiments

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Exp. Num. (from-to)	Number of appointments	Multi-player TAS		Single-player TAS							
		Objective value (unit cost)	Run time (sec)	Objective value (unit cost)				Run time (sec)			
				$C_{arr} = 0.01$	$C_{arr} = 0.1$	$C_{arr} = 0.5$	$C_{arr} = 1$	$C_{arr} = 0.01$	$C_{arr} = 0.1$	$C_{arr} = 0.5$	$C_{arr} = 1$
1-5	4	6,310	2.7	4	30	100.4	150.4	1.8	1.7	1.7	1.8
6-10	6	12,660	10.2	9.0	60	185.4	300.6	1.8	1.8	1.8	1.9
11-15	8	21,950	184.6	16	100	270.6	460.6	1.8	2.0	2.0	1.9
16-20	10	31,180	144.4	25	140	405.6	661.2	2.0	2.1	2.2	2.4
21-25	12	46,499	37.6	34	168.6	528.2	903.2	2.2	2.4	2.2	2.7
26-30	14	61,999	35.3	45	218.6	689.6	1,189.6	2.3	2.4	2.5	2.6
31-35	16	80,600	18.0	56	273.2	872.6	1,537.8	2.4	2.6	3.1	3.3
36-40	18	102,299	25.5	69	331.2	1,055.8	1,895.8	2.5	3.3	3.2	3.2
41-45	20	124,000	139.8	82	389.2	1,287.2	2,290.4	3.0	3.8	3.5	3.7
46-50	22	151,900	748.2	97	459.2	1,523.4	2,748.4	3.3	4.2	4.1	3.9

The average drayage cost of the multi-player model is 55% higher compared to the single-player model with $\mathrm{C_{arr}} = 1$ (gate waiting cost is the only criteria considered in Eq. (3.55)), but it should be noted that the solution of the multi-player model is an equilibrium solution among drayage firms. That is, the solution from the multi-player model is one that accounts for some drayage firms not accepting the assigned appointment and changing their appointments to the next day. In a way, the multi-player model solution is similar to the optimization model (less sensitive to objective function coefficients) in that it accounts for various scenarios. Stated differently, applying the solution from the single-player model (sensitive to objective function coefficients) may result in longer waiting time during certain time windows and lower utilization of gate resources due to drayage firms changing their appointment times or day.

Figure 3.4b shows the effect of problem size on gate waiting cost. As expected, as more emphasis are put on the gate waiting cost component in Equation 3.55 (i.e., higher value for $\mathrm{C_{arr}}$) the lower the gate waiting cost. As shown, the solutions of the single-player model with $\mathrm{C_{arr}} = 0.5$ and 1 are nearly identical to each other and that of the multi-player model. These results suggest that the multi-player model yields the lowest gate waiting cost. In other words, no value of $\mathrm{C_{arr}}$ in the single-player model can produce a solution with lower gate waiting cost.

The summation of drayage cost and gate waiting cost could be as a metric for comparing the single-player and multi-player models. The total cost for the multi-player model averaged across all problem sizes is 234. For the single-player model, the total cost with $\mathrm{C_{arr}}$= 0.01, 0.1, 0.5 and 1 are 667, 301, 230, and 224, respectively. These results confirm what was stated earlier regarding the non-sensitivity of the multi-player model to objective function coefficients.

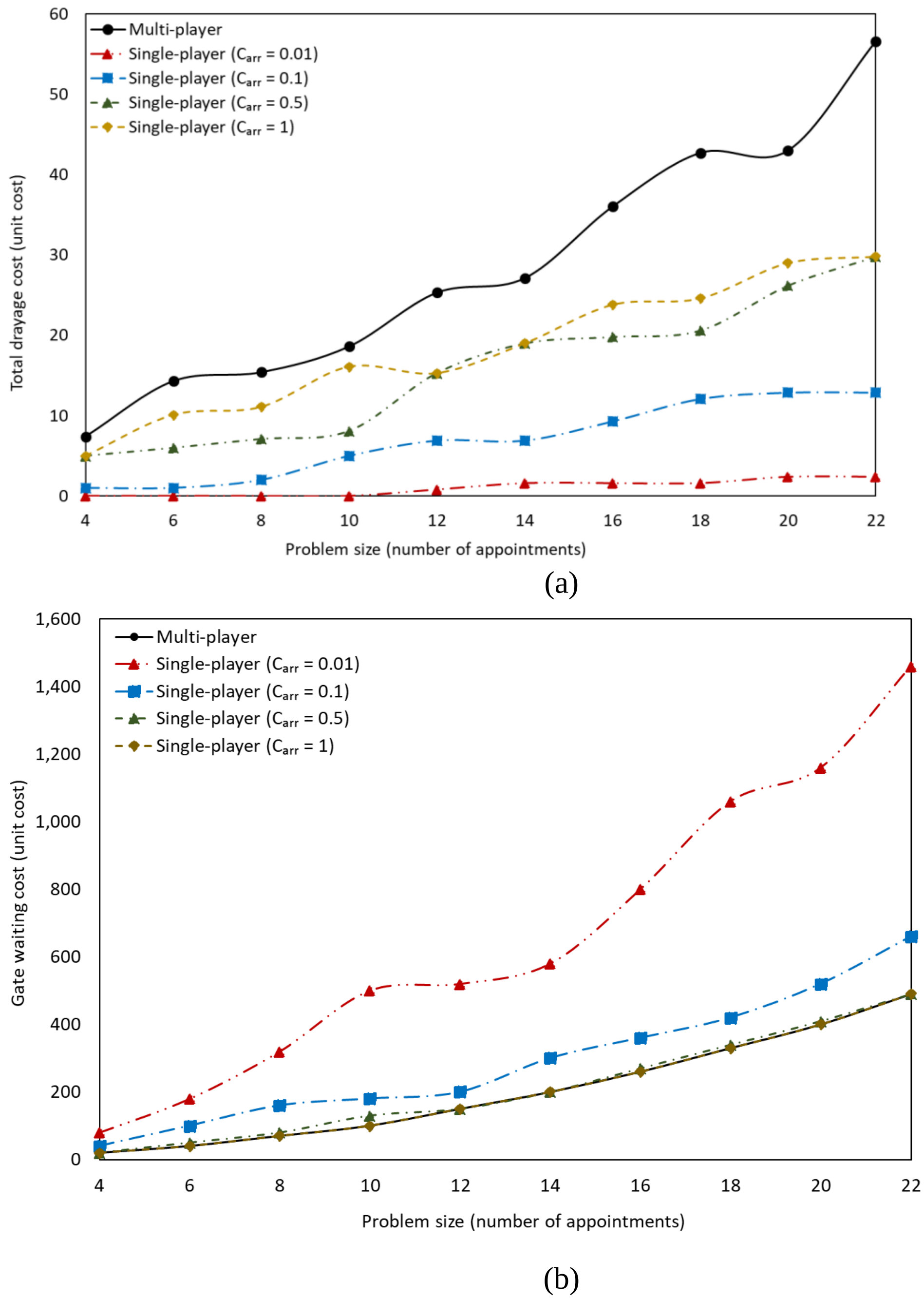


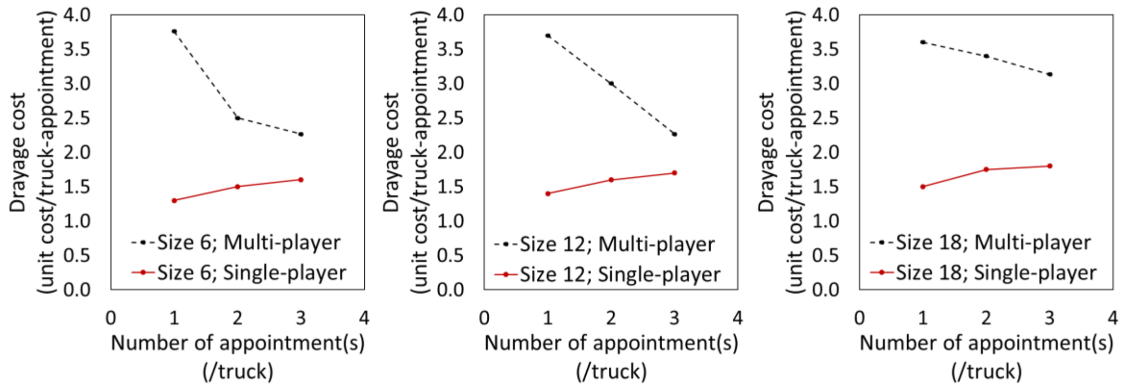
Figure 3.4 Total drayage cost (a) and gate waiting cost (b) as a function of problem size

Table 3.7 shows the objective function value and run time of experiments 51 to 68. This table is separated from Table 3.6 because the single-player model is run for only one scenario with $C_{arr} = 1$. It can be seen that the average run time of the multi-player model increases exponentially with problem size: (2.5+2.9+2.6)/3=2.7, (4.1+9.5+31.8)/3=15.1, and (10.2+17.6+183.2)/3=70.3. The number of appointments per truck also has a negative effect on the run time.

Figure 3.5(a) shows graphically the relationship between drayage cost and number of appointments per truck. It can be seen that for the multi-player model, the higher number of appointments per truck, the higher the drayage cost. Conversely, for the single-player model, the higher number appointments per truck, the lower the drayage cost. This trend is reflected in the results for all three problem sizes. These results suggest that the use of the proposed multi-player model would result in higher cost savings for the drayage firms.

Table 3.7 Results of experiments

(1)	(2)	(3)	(4)	(5)
Exp. Num. (from-to)	Multi-player		Single-player ($C_{arr} = 1$)	
	Objective value (unit cost)	Run (sec)	Objective value (unit cost)	Run (sec)
51-52	125	2.5	265.3	2.4
53-54	125	2.9	270.4	1.9
55-56	125	2.6	270.4	1.8
57-58	465	4.1	917.1	2.2
59-60	465	9.5	931.8	2.2
61-62	465	31.8	942.0	2.1
63-64	1,023	10.2	1,923.6	3.2
65-66	1,023	17.6	1,963.0	3.1
67-68	1,023	183.2	1,974.0	2.8



(a) (b) (c)

Figure 3.5 Drayage cost comparison

3.6 SUMMARY AND CONCLUSIONS

This paper proposed a novel multi-player TAS model using game theory. A bi-level multi-player programming problem is formulated with the marine container terminal function as the leader at the upper-level and multiple drayage firms function as followers at the lower-level. The objective of the leader (the terminal) is to minimize the gate waiting cost of trucks by spreading out the truck arrivals, and the objective of the followers (drayage firms) is to minimize their own drayage cost. To obtain an exact solution, lemmas and linearization techniques are used to transform the bi-level model into a single-level MIP with continuous variable. This study is the first to propose the use of KKT conditions for the lower-level problem to enable the TAS problem to be solved as a standard MIP.

Experimental results indicated that the proposed multi-player model yields a lower gate waiting cost compared to the typically used single-player model. Although its total cost (gate waiting cost + drayage cost) is higher in some cases compared to the single-player model, it has the advantage of being less sensitive to objective function coefficients. That is, its solution takes into account the fact that some drayage firms may not accept the assigned appointment and change their appointments to the next day. In a way, the multi-player model's solution is similar to that of an optimization model's solution which is not sensitive to objective function coefficients. The results also indicated that the use of the proposed multi-player model would result in higher cost savings for the drayage firms as the number of appointments per truck increases. Lastly, the proposed multi-player model has the advantage of being able to be solved in a single run.

The limitations of this study need to be considered when interpreting the results. First, the drayage firms are only allowed to request an appointment for next day. Same day requests may be taken into account in a future study regarding dynamic appointment systems. Second, the drayage firms do not share any information regarding their appointments. So, a follow-up study can investigate the effect of TAS on drayage firm cost while allowing some collaborations among drayage firms.

CHAPTER 4

DESIGN of a TRUCK APPOINTMENT SYSTEM CONSIDERING DRAYAGE SCHEDULING and STOCHASTIC TURN TIME [1]

Truck appointment systems (TAS) are being implemented by more and more marine container terminals in the U.S. to deal with gate congestion since its conception in the early 2000's. Some container terminals in the U.S., such as Port of Baltimore, Port of Virginia, and Port of Vancouver, require trucks to make appointments for their transactions. That is, truckers must notify the terminal operator in advance (typically 24 hours or greater) of their intended transaction (e.g., full or empty container pickup, full or empty container drop-off, chassis pickup or drop-off, dual-transaction, etc.) and select one of the available time-windows (typically 1 hour or longer); trucks without appointments will not be processed and turned away. The Ports of Los Angeles and Long Beach reported that the implementation of the required truck appointment system, along with a weaker import volume, helped lower the turn time to 67 minutes in December 2019 which was the lowest turn time since 2014 (Mongelluzzo, 2020). While a lower turn time is important to terminal operators from an efficiency and environmental perspective, it is even more important to the drayage operators whose productivity relies on the quick processing of

[1] Torkjazi, Mohammad and Nathan Huynh. "Design of a Truck Appointment System Considering Drayage Scheduling and Stochastic Turn Time."

trucks at the terminals in order to complete the rest of their moves before the end of the business day; in this study, turn time is defined as gate-out time minus gate-in time.

A typical TAS allows terminal operators to set quotas for pre-specified time-windows. This is the reason why many studies have sought to design a TAS to optimize the quotas (e.g., Morais and Lord, 2006; Huynh and Walton; 2008; Huynh, 2009; Chen et al., 2011; Chen et al., 2013a; Chen et al., 2013b; Zhang et al, 2013). To determine an appropriate quota, the terminal operator needs to balance the congestion outside its terminal gate and inside the container yard. A low quota will reduce congestion at the gate but will leave the container yard's handling equipment underutilized, and hence, a lower throughput. On the other hand, a high quota will increase the use of the container yard's handling equipment and throughput but will create congestion at the gate. TAS that consider only the terminal's capacity and efficiency may negatively affect the drayage operators. That is, when drayage companies and independent owner operators are unable to come to the terminal at their preferred times, they will have to adjust their schedules which will likely result in suboptimal tours, and therefore, higher cost. More sophisticated TAS address this issue by considering the trucks' preferred times and minimizing the difference between the assigned and preferred times.

A crucial element related to TAS that has not been considered in any prior study, including more sophisticated TAS, is the variability in turn time. This operational characteristic is important to consider because any TAS model that uses a deterministic turn time may provide an appointment schedule that is infeasible for trucks with more than one appointment. Consider the solution illustrated in Figure 4.1 where the appointment system was developed using a deterministic turn time of 30 minutes. In this solution, a truck was informed to arrive at time window 1 (between 9 and 10 AM) and time window 3 (between 11 AM and 12 PM) for its two appointments; in this example, each time window is 1-hour long. Given these time windows, the drayage operator adjusted his schedule such that he will pick up an export container at 8:30 AM and arrive at the terminal by 9:30 AM for his first appointment. With a 30-minute turn time, he will be able to depart the terminal by 10:00 AM and travel to his next destination. Given a total travel time of 60 minutes (out and back) and transaction time of 45 minutes, the drayage operator will be able to make it back to the terminal by 11:45 AM for his second appointment. If due to unforeseen circumstances, the turn time for the first appointment was 60 minutes instead of 30, then the drayage operator will not be able to depart the terminal until 10:30 AM. The earliest time he can make it back to the terminal will be 12:15 PM, which is outside his assigned time-window; he will not receive service if he is late and may even have to pay a penalty fee. This study seeks to address this issue and contribute to the TAS literature by proposing a novel TAS model that explicitly considers the variability in turn time.

The objective of this paper is to develop a new mathematical model for the TAS that seeks to minimize the drayage operation cost. The novelty of this model is that it uses a probability distribution for turn time. The proposed TAS model is formulated as a mixed integer-linear program with chance-constraints to address stochastic turn time. To solve the proposed TAS model, a sample-average approximation (SAA) method is used in conjunction with the CPLEX solver. Using best-fit distributions of actual turn times from a U.S. port, a series of experiments is designed to evaluate the effectiveness of the proposed TAS model as compared to the deterministic model where an average turn time is used.

The rest of the paper is organized as follows. Section 2 provides a summary of literature review on stochastic TAS and drayage models. Section 3 discusses the model formulation and solution methodology. Section 4 discusses the experimental design and results of experiments. Lastly, section 5 provides a summary and concluding remarks.

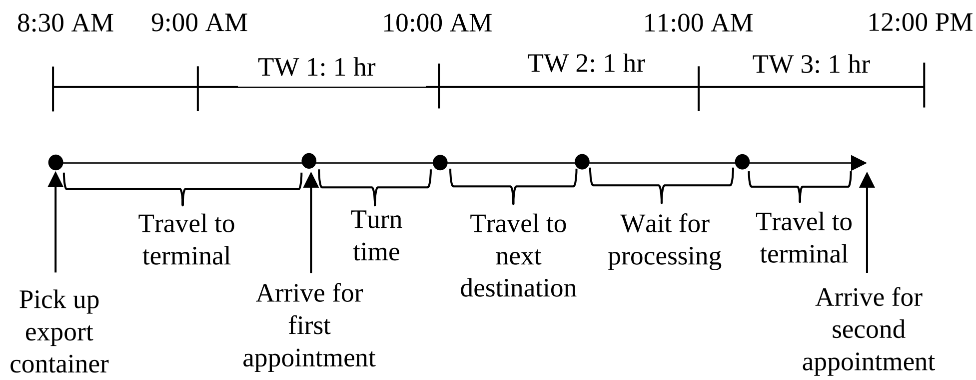


Figure 4.1 Illustration of TAS solution with fixed turn time

4.1 LITERATURE REVIEW

A number of studies have sought to develop methodologies to determine optimal quotas for TAS or assess the effectiveness of TAS. A comprehensive review of TAS can be found in the work of Huynh et al. (2016). The following review focuses on TAS studies that considered uncertainty (i.e., stochasticity) in the durations of certain terminal operation processes.

Li et al. (2018) considered situations when some trucks deviate from their assigned schedules (i.e., not arriving within the assigned time-windows). They proposed a resilience strategy for the TAS to neutralize the impact of late or early arrivals. The authors used discrete event simulation (DES) to evaluate the effectiveness of their proposed strategy. Their simulation results indicated that disruption of external truck arrivals can significantly impact the entire system performance which includes total waiting time of on-time trucks and total idling emissions. Zehendner and Feillet (2014) developed a tool to use the TAS to increase the service quality of trucks, trains, barges, and vessels. They proposed a MILP model to determine the number of appointments to offer with respect to the overall workload and available handling capacity. Their optimization model is based on a network flow representation of the terminal and aims to minimize overall delays at the terminal. It simultaneously determines the number of truck appointments to offer and allocates straddle carriers to different transport modes. They used a DES to validate the findings of their optimization model at an operational level assuming stochastic parameters (arrival times, volumes, and handling times). They compared three operational scenarios to evaluate the impact of a truck appointment system on truck delays using DES: 1) a container terminal operating without an appointment system, 2) an ideal case where the container terminal uses an appointment system and all appointments are respected (ideal appointment system) and 3) a more realistic case where the container terminal uses an appointment system but some trucks do not show up and some trucks are served without an appointment (realistic appointment system). They found that the ideal and realistic appointment systems reduce the average truck service time by 17 to 22 minutes and 14 minutes, respectively. Zhao and Goodchild (2010) evaluated the use of truck arrival information from the TAS to reduce import container rehandling. They found that a significant reduction in rehandles can be obtained by having just a small amount of truck arrival information compared to the case where truck arrivals are stochastic, and no information is provided to the terminal. Chen et al. (2013c) proposed two types of TAS: static TAS and dynamic TAS. The static TAS is an optimization model that seeks to minimize the adjustments between the truckers' preferred arrival times and the assigned appointment time-window, so as to reduce truck rescheduling. For the dynamic TAS, they considered a stochastic truck arrival process; that is, truck arrivals follow a non-homogenous Poisson process and its average arrival rate in a time window is equal to the number of corresponding quotas. Unlike the static TAS that provides an appointment schedule for the next day, the dynamic TAS is a real-time algorithm that allows drayage firms and independent owner operators to make same-day appointments. Individual trucker logs into an online proprietary system to make appointment requests. An appointment request will be evaluated based on the existing appointments: if the appointment request will not result in a long queue or a long waiting time for trucks, then it will be accepted; otherwise, the request will be rejected and the trucker will have to make a new request. Estimated queuing times in each period of the day based on existing appointments are provided to assist users in selecting appointment time-windows. Through comparison of their proposed TAS models, they concluded that the dynamic TAS has several advantages over the static TAS: 1) it can provide same-day appointment requests, and 2) it provides more appointment choices. Azab et al. (2019) considered stochastic gate service time, inter-terminal travel time, container yard crane handling rate, quay crane handling rate, and failure of equipment in their developed DES model. They found that their proposed TAS DES-based model can reduce the turn time by 29% and queuing time at the gate by 38%.

A number of studies have sought to consider the impact of TAS on drayage scheduling. Only those drayage scheduling studies that considered uncertainty for some aspects of drayage operations are included in this review. Máhr et. al (2010) considered uncertainty in customer and terminal service times and job arrivals. They developed two approaches, an agent-based model and an online optimization model for drayage problem with TAS constraint. They found that the agent-based model outperforms the optimization model when service time is highly uncertain. Marković et. al (2014) considered uncertainty in truck round trip durations and train departure times while solving the drayage problem with terminal time window constraint. Their proposed drayage model minimizes the storage cost, in-terminal operation cost and late delivery penalty cost. They proposed two solution methodologies, a local search heuristic based on interior point method and a hybrid genetic algorithm to solve the model. They found that the drayage cost can be reduced by increasing the available storage capacity and allowed number of trucks to arrive per time window. Shiri and Huynh (2019) proposed an integrated drayage scheduling model that accounts for the uncertainty in container loading and unloading times. They employed a Tabu Search algorithm to solve their model. Their numerical experiment results indicated that their proposed model produce schedules that are more likely to be feasible under a variety of scenarios compared to the deterministic model.

In this study, we propose a TAS optimization model that considers stochastic turn times at marine container terminals. As indicated in Table 4.1, no prior study has included this operational characteristic in its optimization model and investigated how it affects TAS. It builds on our previous work (Torkjazi et al. (2018)) which proposed a deterministic TAS that explicitly considered drayage truck tours. That is, the TAS took as input the drayage tours of trucks and create an assignment schedule that minimized the impact to drayage tours. A nonlinear fluid-based approximation method was used to calculate the queue length at the gate of the terminal, and hence, the proposed TAS was formulated as a mixed integer nonlinear program (MINLP) and the model was solved using built-in nonlinear solvers included in the Lingo commercial software. The current work differs from our previous work in that it 1) uses a linear function to estimate the queue length which allows the TAS model to be formulated as a MILP and solvable using CPLEX, 2) replaces the deterministic turn time constraint with a chance-constraint, 3) replaces the objective function terms with their expected values which transforms the deterministic TAS model into a stochastic one, and 4) utilizes a SAA algorithm to solve the stochastic MILP model.

Figure 4.2 shows the TAS framework utilized in this study. In this framework, the TAS serves one maritime container terminal and multiple drayage firms. Each drayage firm owns a number of trucks that can be used to perform container move requests, and each request must be performed within a pre-defined customer's time-window and assigned appointment time-window at the maritime container terminal. Based on the Port of Vancouver TAS model, it is assumed that on a daily basis drayage firms will have submitted their appointment requests by 5:00 PM for the next day's appointments (step 1). These appointment requests correspond to the times when trucks need to be at the terminal based on the optimal tours. In this study, the optimal truck tours are determined using a drayage scheduling model, based on the work of Shiri and Huynh (2016). For each appointment request, the information to be provided includes truck ID and container number. Similarly, it is assumed that the terminal operator will have submitted the quotas for each time-window and most recent turn time distributions for export, import, and dual transactions by 5 PM (step 2).

Table 4.17 Summary of stochastic TAS and drayage scheduling literature review

Authors	TAS	DR	Solution method		Stochastic parameters								
			Sim	Opt	TT	TAT	HT	GST	ITT	QT	DTD	CST	TRT
Li et al. (2018)	✓		✓			✓							
Zehendner, Feillet (2014)	✓		✓			✓	✓						
Zhao, Goodchild (2010)	✓		✓				✓						
Chen et al. (2013c)	✓			✓		✓							
Azab et al. (2019)	✓		✓				✓	✓	✓				
Chen et al. (2013a)	✓			✓						✓			
Máhr et. al (2010)		✓	✓	✓	✓							✓	
Marković et. al (2014)		✓		✓									✓
Shiri, Huynh (2019)		✓		✓							✓	✓	
This study	✓	✓		✓	✓								

*TAS: TAS with stochastic parameters; DR: Drayage scheduling with stochastic parameters and TAS constraint; Sim: Simulation; Opt: Optimization; TT: Turn time; TAT: Truck arrival time; HT: Handling time; GST: Gate service time; ITT: Inter-terminal travel time; QT: Queuing time; DTD: Drayage task duration; CST: Customer service time; TRT: Truck round trip.

4.2 PROBLEM DESCRIPTION AND FORMULATION

Based on all the appointment requests associated with a truck ID, the TAS will determine the respective "desired" time-gaps between consecutive appointments. It is assumed that a truck will seek to perform as many container moves as possible in a given day, and thus, the time-gap between requested appointments represents the absolute minimum time the truck needs to complete the job; note that this time-gap includes the truck turn time. Once all of the input data (appointment requests and quotas) are provided, the TAS then determines the appointment time-window for each request with the objective of minimizing the cost of changing the time-gap between two consecutive appointments of a truck with more than one appointment, cost of changing an appointment to another time window, and cost of queuing at the gate due to congestion (step 3). Lastly, the optimal appointments for each truck are sent to the drayage firms (step 4). The assigned appointment time-windows may require drayage firms to reschedule their trucks' tours if they are different from the requested times. The economic impact of TAS on drayage firms can be determined by computing the difference between the total drayage cost of the original tours and the cost of the adjusted tours. The model that provides the original tours is referred to as the "unrestricted drayage model" hereafter because this model obtained the optimal truck tours without any restriction on when they need to be at the terminal. The model that provides the adjusted tours is referred to as the "restricted drayage model" because this model obtained the optimal truck tours with restriction on when they need to be at the terminal.

When formulating the TAS, it is assumed that all trucks are required to make an appointment as it is in practice at some terminals (e.g., TraPac terminal at port of Oakland, Evergreen terminal at port of Los Angeles). It is also assumed that the container terminal will process trucks during the appointment time-window and that no appointment will be transferred to the next time-window because of the lack of resources at the container terminal. In practice, trucks are able to pick up and deliver containers before and after terminal hours; thus, it is assumed that a truck can pick up the container from the customer up to 4 h before the terminal opens and deliver the container to the customer up to 4 h after the terminal closes. It is also assumed that drayage firms do not have knowledge about other requested appointments and the terminal-specified quotas.

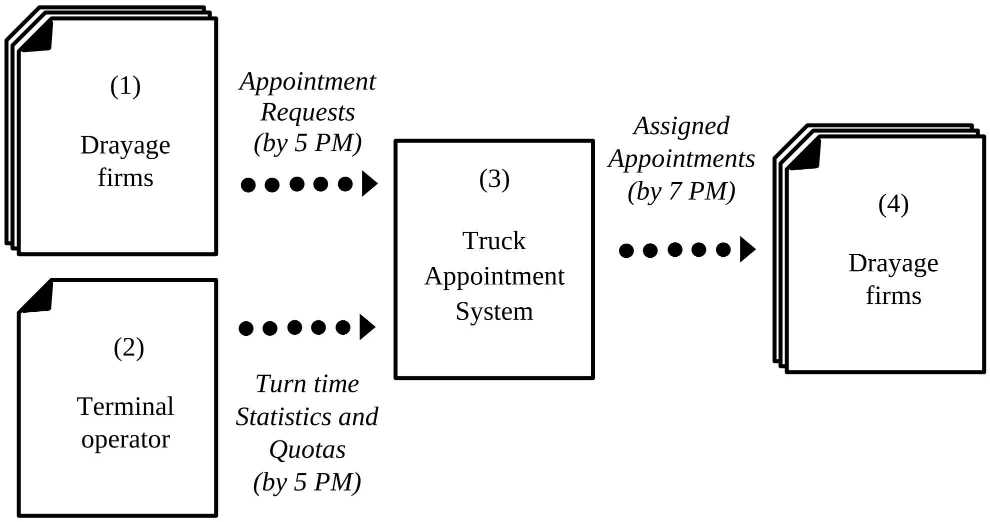


Figure 4.2 Illustration of TAS framework

The formulation considers three separate cases for turn times since their processes and durations are distinctly different in practice: export, import and dual (export and import transactions are performed in a single trip). Given that turn times are probability distributions rather than constants, the constraints (see constraints (4) to (6), (9) to (11), (15) to (17), and (21) to (23) below) which calculate the difference between the assigned time-windows and the preferred time-windows are formulated as chance-constraints; the reason is because the preferred and assigned time-windows for trucks with more than one appointment implicitly include turn times of previous appointments (except for the first appointment). Specifically, these constraints subtract the fixed turn time value used by the drayage firm to determine the tour and add a turn time distribution received from the terminal operator.

Sets, subsets, indices, parameters and decision variables of the stochastic TAS model are shown in Table 4.2.

Table 4.2 Sets, subsets, indices, parameters and decision variables of TAS models

Sets		*Subsets*		*Indices*	
D	Drayage firms	K_d	Trucks from drayage firm d	t	Time-window
K	Trucks	A_i	Number of appointments for trucks with i appointments	j	Extra arrival
A	Appointment numbers			k	Truck
T	Time-windows			a	Appointment
J	Extra arrivals			d	Drayage firm
N_I	Import appointments			n	Sample size
N_E	Export appointments				
N	Sample sizes				

Parameters	
p_{ka}	Preferred arrival time of a^{th}appointment of truck k
C_{large}^{gap}	Penalty value for actual time-gap larger than the preferred time-gap
C_{small}^{gap}	Penalty value for actual time-gap smaller than the preferred time-gap
C_{pos}	Penalty value applied to truck that arrives earlier than scheduled
C_{neg}	Penalty value applied to truck that arrives later than scheduled
C_{arr}	Penalty value applied to length of queue at the gate
M	Auxiliary gate waiting cost effect (unit cost per gate waiting cost)
C_t	Quota at time-window t
$\tilde{\omega}_E^n$	Stochastic turn time for export transaction
$\tilde{\omega}_I^n$	Stochastic turn time for import transaction
$\tilde{\omega}_{EI}^n$	Stochastic turn time for dual transaction
ω^p	Deterministic turn time used by drayage firms for drayage scheduling problem
n_E	Number of export transaction subtasks
n_I	Number of import transaction subtasks
n_{EI}	Number of dual transaction subtasks

Table 4.2 Sets, subsets, indices, parameters and decision variables of TAS models (Continued)

Decision variables	
X_{ka}	Arrival time of a^{th} appointment of truck k
N_{ka}^{n}	Difference between time-gap between a^{th} and $(a+1)^{th}$ appointments of truck k
Q_{ka}^{n}	Difference between time-gap between $(a+1)^{th}$ and a^{th} appointments of truck k
S_{ka}^{n}	Difference between actual and preferred arrival time-window
Z_{ka}^{n}	Difference between preferred and actual arrival time-window
D_{tj}	1 if time-window t has more than j arrivals, 0 otherwise
Y_{t}	Number of arrivals at time-window t
U_{tj}	Penalty value for j^{th} arrivals at time-window t
V_{t}	Gate waiting cost at time-window t

$$\text{Min } E_c\left[C_{large}^{gap}\sum_{k\in K_d}\sum_{a\in A_i}N_{ka}^{n}\right]+E_c\left[C_{small}^{gap}\sum_{k\in K_d}\sum_{a\in A_i}Q_{ka}^{n}\right]+E_c\left[C_{pos}\sum_{k\in K_d}\sum_{a\in A_i}S_{ka}^{n}\right]+E_c\left[C_{neg}\sum_{k\in K_d}\sum_{a\in A_i}Z_{ka}^{n}\right]+C_{arr}\sum_{t\in T}V_t \quad (4.1)$$

Subject to:

$$N_{ka}^{n}\geq 0 \qquad \forall k\in K_d, \forall a\in A_i\backslash\{i\}: i>1, n\in N \quad (4.2)$$

$$\sum_{t\in T}t\cdot X_{k(a+1)t}-\sum_{t\in T}t\cdot X_{kat}-\left(p_{k(a+1)}-p_{ka}\right)\leq N_{ka}^{n} \qquad \forall k\in K_d, \forall a\in A_i\backslash\{i\}: i>1, p_{ka}=p_{k(a+1)}, n\in N \quad (4.3)$$

$$\sum_{t\in T}t\cdot X_{k(a+1)t}-\sum_{t\in T}t\cdot X_{kat}-\left(p_{k(a+1)}-p_{ka}+\widetilde{\omega}_{E}^{n}-\omega^{p}\right)\leq N_{ka}^{n} \qquad \forall k\in K_d, \forall a\in A_i\backslash\{i\}: i>1, p_{ka}\neq p_{k(a+1)}, a\in N_E, n\in N \quad (4.4)$$

$$\sum_{t\in T}t\cdot X_{k(a+1)t}-\sum_{t\in T}t\cdot X_{kat}-\left(p_{k(a+1)}-p_{ka}+\widetilde{\omega}_{I}^{n}-\omega^{p}\right)\leq N_{ka}^{n} \qquad \forall k\in K_d, \forall a\in A_i\backslash\{i\}: i>1, p_{ka}\neq p_{k(a+1)}, a\in N_I, n\in N \quad (4.5)$$

$$\sum_{t\in T}t\cdot X_{k(a+1)t}-\sum_{t\in T}t\cdot X_{kat}-\left(p_{k(a+1)}-p_{ka}+\widetilde{\omega}_{EI}^{n}-\omega^{p}\right)\leq N_{ka}^{n} \qquad \forall k\in K_d, \forall a\in A_i\backslash\{i\}: i>2, p_{ka}\neq p_{k(a+1)}, a\in N_{EI}, n\in N \quad (4.6)$$

$$Q_{ka}^{n}\geq 0 \qquad \forall k\in K_d, \forall a\in A_i\backslash\{i\}: i>1, n\in N \quad (4.7)$$

$$p_{k(a+1)} - p_{ka} - \left(\sum_{t\in T} t \cdot X_{kat} - \sum_{t\in T} t \cdot X_{k(a+1)t}\right) \leq Q_{ka}^{n} \quad \forall k \in K_d, \forall a \in A_i \backslash \{i\}: i > 1; p_{ka} = p_{k(a+1)}, n \in N \tag{4.8}$$

$$p_{k(a+1)} - p_{ka} + \widetilde{\omega}_{E}^{n} - \omega^{p} - \left(\sum_{t\in T} t \cdot X_{kat} - \sum_{t\in T} t \cdot X_{k(a+1)t}\right) \leq Q_{ka}^{n} \quad \forall k \in K_d, \forall a \in A_i \backslash \{i\}: i > 1; p_{ka} \neq p_{k(a+1)}; a \in N_E, n \in N \tag{4.9}$$

$$p_{k(a+1)} - p_{ka} + \widetilde{\omega}_{I}^{n} - \omega^{p} - \left(\sum_{t\in T} t \cdot X_{kat} - \sum_{t\in T} t \cdot X_{k(a+1)t}\right) \leq Q_{ka}^{n} \quad \forall k \in K_d, \forall a \in A_i \backslash \{i\}: i > 1; p_{ka} \neq p_{k(a+1)}; a \in N_I, n \in N \tag{4.10}$$

$$p_{k(a+1)} - p_{ka} + \widetilde{\omega}_{EI}^{n} - \omega^{p} - \left(\sum_{t\in T} t \cdot X_{kat} - \sum_{t\in T} t \cdot X_{k(a+1)t}\right) \leq Q_{ka}^{n} \quad \forall k \in K_d, \forall a \in A_i \backslash \{i\}: i > 2; p_{ka} \neq p_{k(a+1)}; a \in N_{EI}, n \in N \tag{4.11}$$

$$S_{ka}^{n} \geq 0 \quad \forall k \in K_d, \forall a \in A_i, n \in N \tag{4.12}$$

$$\sum_{t\in T} t \cdot X_{kat} - p_{ka} \leq S_{ka}^{n} \quad \forall k \in K_d, \forall a \in A_i, a = 1, n \in N \tag{4.13}$$

$$\sum_{t\in T} t \cdot X_{kat} - p_{ka} \leq S_{ka}^{n} \quad \forall k \in K_d, \forall a \in A_i, a > 1, p_{ka} = p_{k(a-1)}, n \in N \tag{4.14}$$

$$\sum_{t\in T} t \cdot X_{kat} - (p_{ka} + \widetilde{\omega}_{E}^{n} - \omega^{p}) \leq S_{ka}^{n} \quad \forall k \in K_d, \forall a \in A_i, a > 1, p_{ka} \neq p_{k(a-1)}, (a-1) \in N_E, n \in N \tag{4.15}$$

$$\sum_{t\in T} t \cdot X_{kat} - (p_{ka} + \widetilde{\omega}_{I}^{n} - \omega^{p}) \leq S_{ka}^{n} \quad \forall k \in K_d, \forall a \in A_i, a > 1, p_{ka} \neq p_{k(a-1)}, (a-1) \in N_I, n \in N \tag{4.16}$$

$$\sum_{t\in T} t \cdot X_{kat} - (p_{ka} + \widetilde{\omega}_{EI}^{n} - \omega^{p}) \leq S_{ka}^{n} \quad \forall k \in K_d, \forall a \in A_i, a > 2, p_{ka} \neq p_{k(a-1)}, (a-1) \in N_{EI}, n \in N \tag{4.17}$$

$$Z_{ka}^{n} \geq 0 \quad \forall k \in K_d, \forall a \in A_i, n \in N \tag{4.18}$$

$$p_{ka} - \sum_{t\in T} t \cdot X_{kat} \leq Z_{ka}^{n} \quad \forall k \in K_d, \forall a \in A_i, a = 1, n \in N \tag{4.19}$$

$$p_{ka} - \sum_{t\in T} t \cdot X_{kat} \leq Z_{ka}^{n} \quad \forall k \in K_d, \forall a \in A_i, a > 1, p_{ka} = p_{k(a+1)}, n \in N \tag{4.20}$$

$$(p_{ka} + \widetilde{\omega}_{E}^{n} - \omega^{p}) - \sum_{t\in T} t \cdot X_{kat} \leq Z_{ka}^{n} \quad \forall k \in K_d, \forall a \in A_i, a > 1, p_{ka} \neq p_{k(a-1)}, (a-1) \in N_E, n \in N \tag{4.21}$$

$$(p_{ka} + \widetilde{\omega}_{I}^{n} - \omega^{p}) - \sum_{t\in T} t \cdot X_{kat} \leq Z_{ka}^{n} \quad \forall k \in K_d, \forall a \in A_i, a > 1, p_{ka} \neq p_{k(a-1)}, (a-1) \in N_I, n \in N \tag{4.22}$$

$$(p_{ka} + \widetilde{\omega}_{EI}^{n} - \omega^{p}) - \sum_{t \in T} t \cdot X_{kat} \leq Z_{ka}^{n} \qquad \forall k \in K_d, \forall a \in A_i, a > 2, p_{ka} \neq p_{k(a-1)}, (a-1) \in N_{EI}, n \in N \tag{4.23}$$

$$\sum_{t \in T} t \cdot X_{kat} \leq \sum_{t \in T} t \cdot X_{k(a+1)t} \qquad \forall k \in K_d, \forall a \in A_i \tag{4.24}$$

$$\sum_{t \in T} t \cdot X_{kat} = 1 \qquad \forall k \in K_d, \forall a \in A_i, \forall t \in T \tag{4.25}$$

$$\sum_{k \in K_d} \sum_{a \in A_i} X_{kat} \leq C_t \qquad \forall t \in T \tag{4.26}$$

$$Y_t = \sum_{k \in K_d} \sum_{a \in A_i} X_{kat} \qquad \forall t \in T \tag{4.27}$$

$$Y_{t,j} \leq M\left(D_{t,j}\right) + j - 1 \qquad \forall t \in T, \forall j \in J \tag{4.28}$$

$$\sum_{j \in J} U_{tj} \cdot D_{tj} \leq V_t \qquad \forall t \in T \tag{4.29}$$

$$D_{t,j} \in \{0,1\} \qquad \forall t \in T, \forall j \in J \tag{4.30}$$

$$X_{kat} \in \{0,1\} \qquad \forall k \in K_i, \forall a \in A_i, \forall t \in T \tag{4.31}$$

Eq. (4.1) is the objective function that seeks to minimize the total cost. The first and second terms are the expected value of costs associated with increasing and decreasing the gap between two consecutive assigned appointments of a truck compared to those of preferred appointments. The third and fourth terms are the expected value of costs associated with shifting a preferred appointment to an earlier or later time-window. The five terms of the objective function are assigned different penalty values ($C_{large}^{gap}, C_{small}^{gap}, C_{pos}, C_{neg}$, and C_{arr}) to allow the different costs to be weighted differently. The fifth term is the cost of queuing time at the gate of the terminal.

Constraints (4.2) to (4.6) are the linearized forms of the maximum of zero and gap difference (actual gap minus preferred gap) for two consecutive appointments of a truck with more than one appointment. This set of constraints are for the cases when the actual time-gap is greater than the preferred time-gap. Constraint (4.2) is the non-negativity constraint for the auxiliary variable used for linearization. Constraints (4.3) to (4.6) ensure the difference between actual and preferred gaps is less than the auxiliary variable. Constraint (4.3) accounts for two appointments involved in a dual transaction and since the second appointment of dual transactions is not affected by the turn time of the first transaction, this constraint is not a chance-constraint. It should be noted that drayage scheduling model and TAS model consider two appointments for a dual transaction even though both appointments are processed together at gate entry. Constraints (4.4) to (4.6) are chance-constraints associated with export, import, and dual transactions, respectively, for trucks with more than one appointment. As explained previously, the time difference between two consecutive appointments implicitly include the turn time of the first appointment. Due to turn time being stochastic, these constraints are formulated as chance-constraints.

Constraints (4.7) to (4.11) are the linearized forms of the maximum of zero and gap difference (preferred gap minus actual gap) for two consecutive appointments of a truck with more than one appointment. This set of constraints are for the cases when the actual time-gap is less than the preferred time-gap. Constraint (4.7) is the non-negativity constraint for the auxiliary variable used for linearization. Constraints (4.8) to (4.11) ensure the difference between actual and preferred gaps is less than the auxiliary variable. Constraint (4.8) accounts for two appointments involved in a dual transaction and since the second appointment of dual transactions is not affected by the turn time of the first transaction, this constraint is not formulated a chance-constraint. Constraints (4.9) to (4.11) are chance-constraints associated with export, import, and dual transactions, respectively, for trucks with more than one appointment.

Constraints (4.12) to (4.17) are the linearized forms of the maximum of zero and time window difference between actual and preferred time-windows of a truck. This set of constraints are for those cases when the assigned time-window is later than the preferred time window. Constraint (4.12) is the non-negativity constraint for the auxiliary variable used for linearization. Constraints (4.13) to (4.17) ensure the difference between actual and preferred time windows is less than the auxiliary variable. Constraint (4.13) accounts for the first appointment of every truck and since the arrival time for the first appointment is not affected by the turn time of an earlier appointment, this constraint does not include a stochastic turn time, and hence, it is not formulated as a chance-constraint. Constraint (4.14) accounts for the second appointment of a dual transaction and since the second appointment of dual transactions is not affected by the turn time of the first transaction, this constraint is not formulated as a chance-constraint. Constraints (4.15) to (4.17) are chance-constraints associated with export, import, and dual transactions, respectively, for trucks with more than one appointment. Constraints (4.18) to (4.23) are nearly identical to constraints (4.12) to (4.17); the only difference is that they account for situations when the assigned time-windows are earlier than the preferred time windows.

Constraint (4.24) maintains the order of consecutive appointments. Constraint (4.25) ensures that all appointments are met. Constraint (4.26) is the quota constraint. Eq. (4.27) calculates the number of arrivals in every time-window. Constraint (4.28) penalizes each appointment based on number of arrivals in that time-window. Constraint (4.29) calculates the gate waiting cost at each time window to be minimized in the objective function. Lastly, constraints (4.30) and (4.31) define the domain of the decision variables.

4.3 SOLUTION METHODOLOGY

As mentioned, the unrestricted and restricted drayage scheduling models used in this study are based on the work of Shiri and Huynh (2016). They are used to determine the optimal tours from which the arrival times of each truck to the terminal are obtained. The drayage scheduling model is solved using CPLEX and Python. The arrival time windows obtained from the drayage models are input to the deterministic and stochastic TAS models. The deterministic TAS model consists of Eq. (4.1) as objective function and constraints (4.2) to (4.31). The differences between deterministic TAS and stochastic TAS are that 1) the expected values are removed from the objective function, 2) the stochastic turn times ($\tilde{\omega}_{E}^{n}$, $\tilde{\omega}_{I}^{n}$, and $\tilde{\omega}_{EI}^{n}$) are replaced with the average turn times (ω_{E}, ω_{I}, and ω_{I}), and 3) the sample size index *n* is removed from associated decision variables (N_{ka}^{n}, Q_{ka}^{n}, S_{ka}^{n}, and Z_{ka}^{n}). The static TAS model is used as a benchmark for evaluating the effectiveness of the stochastic TAS model. Since the TAS model is a MILP, it can be solved using CPLEX. For the stochastic TAS model, because of the chance-constraints and the expectation term in the objective function, it cannot be solved using CPLEX. In this study, the sample averaging approximation algorithm (SAA) developed by Santoso et al. (2005) is used to estimate the objective function value. Essentially, the SAA draws a random turn time from the given distribution and solves the TAS model (Equations 4.1 to 4.31) using CPLEX. It does this multiple times. The essence of the SAA algorithm is that it guarantees that the final solution is an unbiased estimator of the expected cost expressed in the objective function.

SAA Algorithm

Step 1. Generate M independent samples, each of size N (n = 1, . . ., N), i.e., $(\omega_j^1, \dots, \omega_j^N)$ for j = 1, . . ., M. For each sample, solve the corresponding problem using CPLEX. The SAA objective function ($\hat{\mu}_N$) is calculated using Eq. (4.32).

$$\hat{\mu}_N = \min \frac{1}{N} \sum_{n=1}^{N} T\,(x, \omega_j^n) \tag{4.32}$$

Where $T(x, \omega_j^n)$ represents the objective function of sample j with size n = 1, . . ., N, ω_j^n represents the sample j turn time with size n = 1, . . ., N, and x represents the decision variables.

Step 2. Let μ_N^j and $\widehat{\psi}_N^j$ be the optimal objective function value and optimal solution corresponding to sample j. Compute the average objective function value ($\bar{\mu}_N$) and average variance ($\sigma^2_{\bar{\mu}_N}$) using Eq. (4.33) and (4.34):

$$\bar{\mu}_{N,M} \coloneqq \frac{1}{M} \sum_{j=1}^{M} \mu_N^j \tag{4.33}$$

$$\sigma^2_{\bar{\mu}_{N,M}} \coloneqq \frac{1}{M(M-1)} \sum_{j=1}^{M} (\mu_N^j - \bar{\mu}_{N,M})^2 \tag{4.34}$$

$\bar{\mu}_{N,M}$ is a lower bound to the optimal objective function value of the true problem (μ^*). $\sigma^2_{\bar{\mu}_{N,M}}$ is an estimate of the variance of $\bar{\mu}_{N,M}$. It should be noted that the objective function used for the SAA algorithm, Eq. (4.32), is an unbiased ($E[\widehat{T}_N(x)] = T(x)$) and consistent ($\widehat{T}_N(x) \longrightarrow T(x)$) estimator with probability of 1 as $N \longrightarrow \infty$ *(21)*.

Step 3. For every unique optimal solution from M samples ($\widehat{\psi}_N^j$), fix the optimal solution in the model, use another N' sample size $(\omega^1, \dots, \omega^{N'})$, and calculate the objective function value and variance of the model using Eq. (4.35) and (4.36). The value of N' should be higher than N.

$$\bar{\mu}_{N'}(\bar{x}) \coloneqq \frac{1}{N'}\sum_{n=1}^{N'} T(\bar{x}, \omega^n) \tag{4.35}$$

$$\sigma^2_{N'}(x) \coloneqq \frac{1}{N'(N'-1)}\sum_{n=1}^{N'}(T(\bar{x}, \omega^n) - \bar{\mu}_{N'}(\bar{x}))^2 \tag{4.36}$$

Step 4. Calculate the optimality gap ($gap_{N,M,N'}(\bar{x})$), and the variance of the gap (σ^2_{gap}), using Eq. (4.37) and (4.38):

$$gap_{N,M,N'}(\bar{x}) \coloneqq \bar{\mu}_{N'}(\bar{x}) - \bar{\mu}_{N,M} \tag{4.37}$$

$$\sigma^2_{gap} = \sigma^2_{N'}(x) + \sigma^2_{\bar{\mu}_{N,M}} \tag{4.38}$$

The SAA has two types of errors: bias and variability (Birge and Louveaux, 2011) Bias can be reduced by increasing N, and variability can be reduced by increasing N, or M, or both (Birge and Louveaux, 2011). Increasing either M or N makes the model harder to solve (higher M increases the number of times the model needs to be solved, and higher N increases the number of constraints in the model) and increases the computation time. An experiment was performed where M was set to 10 and N is increased until the optimality gap is equal to or less than 0.01. Generally, a higher N value is required for larger-sized problems. This value of N was found to be different for each different problem size.

4.4 EXPERIMENTAL DESIGN

A series of experiments was performed to investigate the impact of turn time variability on drayage cost. That is, the total cost incurred by the drayage companies as a result of following the appointment schedule produced by the stochastic TAS model is compared against the total cost incurred if they were to follow the schedule produced by the deterministic TAS model. Note that the total drayage cost is obtained from using the "restricted" drayage scheduling model because trucks need to be at the terminal at the assigned time-windows. When evaluating the effectiveness of the stochastic and deterministic TAS schedules, 30 sets of random turn times (export, import, and dual) are used. Each of these 30 independent scenarios represent the situation that may happen during the next day at the container terminal. From these, an average "drayage cost error" is calculated as shown in Figure 4.3. The "drayage cost error" provides a measure of how much the appointment schedule affected drayage operations; the lower this value is, the better it is for drayage firms. The experiments were aimed to investigate the effect of the number of appointments and quotas on drayage cost. To this end, the appointments were varied between 6 to 108 appointments, and the ratio of quotas to number of requested appointments was varied between 1.5 and 3.5.

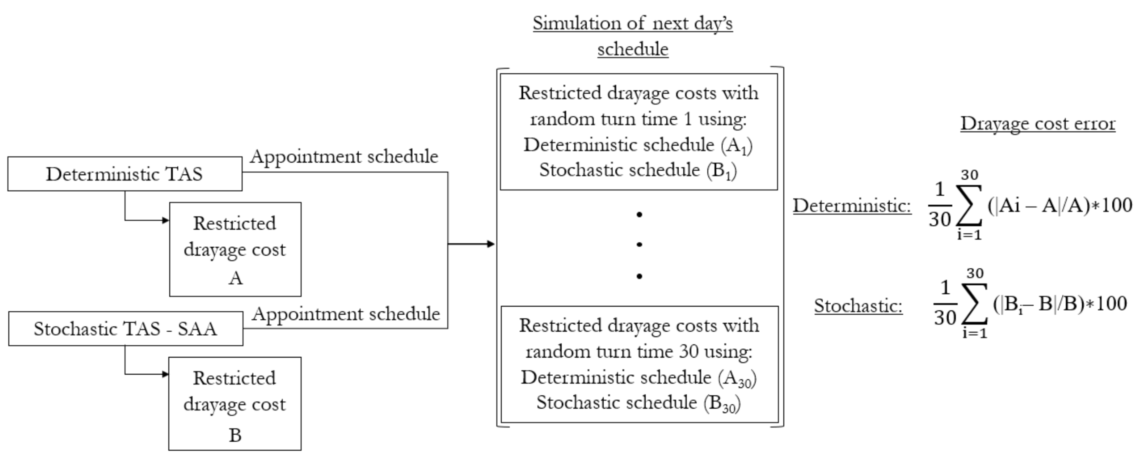


Figure 4.3 Process of calculating drayage cost error

Experiments were generated on a hypothetical square network with deterministic travel times of 3 hours on each edge of the square (Torkjazi et al., 2018). In this network, a truck can perform up to three appointments per day. The customer locations were assumed to be scattered randomly throughout the network. The location of marine container terminal was assumed to be at the mid-point of the west edge of the network. The pickup and delivery time-window of the customers was assumed to be between 4:00 AM to 10:00 PM and the terminal was assumed to operate from 8:00 AM to 6:00 PM. The terminal was assumed to have ten time-windows and each one is one-hour long. The penalty parameters , $C_{large}^{gap}, C_{small}^{gap}, C_{pos}, C_{neg}$, and C_{arr} were assumed to be 1, 3, 1, 3, and 1, respectively (Torkjazi et al., 2018). The turn time distributions (export, import, and dual) used where were generated from data at a U.S. container terminal. The distribution parameters are provided as 1) Export: Wakeby distribution $(a = 3.697, b = 25.929, g = 0.48173, d = -0.55833, x = 0.18288)$, 2) Import: Wakeby distribution $(a = 1.0418, b = 11.05, g = 0.57873, d = -0.59493, x = 0.31784)$, and 3) Dual: Wakeby distribution $(a = -0.49948, b = 2.0962, g = 0.72623, d = -0.67052, x = 0.84476)$. The experiments were performed on a desktop computer with Intel Core i7 3.4 GHz CPU and 16 GB of RAM.

4.5 RESULTS AND DISCUSSION

Table 4.3 (Columns 1 to 5) shows the results of the stochastic TAS model using SAA algorithm. Column 1 shows the experiment number. Column 2 shows the problem size in terms of number of appointments. Column 3 shows the chosen value for N as discussed in the Solution Methodology section; the chosen value is the smallest value that yields an optimality gap ≤ 0.01. Column 4 provides the objective function value for the corresponding problem size. Column 5 shows the optimality gap of the SAA solution. By design, M and N were chosen to ensure that the optimality gap is always ≤ 0.01. A low optimality gap indicates that the objective function value of the optimal solution is sufficiently close to the average objective function value calculated using different random turn times. It can be observed that the objective function value increases linearly with problem size, at a rate of about 500 unit-costs per appointment.

Table 4.3 Results of stochastic TAS model and drayage cost for different problem sizes

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
		Stochastic TAS Model			Drayage cost (unit cost per appointment)			Drayage cost error (%)		
Experiment number	Problem size (number of appointments)	N	Objective function value (unit cost)	Optimality gap	Deterministic TAS	Stochastic TAS	Difference	Deterministic TAS	Stochastic SAA TAS	Difference
1	6	1	113.7	0.010	296.33	307.17	10.83	10.0	5.1	-4.9
2	12	3	206.2	0.002	306.83	313.67	6.83	12.6	10.2	-2.4
3	18	7	339.3	0.006	308.50	314.06	5.56	10.9	8.9	-2.0
4	24	5	506.1	0.001	277.88	283.46	5.58	16.9	11.9	-5.0
5	30	1	827.4	0.006	257.13	264.00	6.87	15.1	10.0	-5.1
6	36	1	1715.5	0.0005	313.78	318.42	4.64	11.3	8.6	-2.7
7	42	5	1958.4	0.001	265.38	275.24	9.86	13.1	9.2	-3.9
8	48	3	2111.4	0.008	268.44	275.52	7.08	12.1	8.6	-3.5
9	54	7	2430.8	0.002	291.43	292.85	1.43	15.1	11.6	-3.5
10	60	5	3391.0	0.008	269.60	275.33	5.73	13.7	8.8	-4.9
11	66	9	3942.2	0.001	271.97	283.86	11.89	12.4	7.5	-4.9
12	72	9	4830.7	0.005	273.43	282.82	9.39	12.1	8.6	-3.5
13	78	11	5039.5	0.005	275.64	289.87	14.23	16.0	9.6	-6.4
14	84	13	5480.8	0.002	265.69	271.08	5.39	12.8	9.8	-3.0
15	90	15	5877.2	0.001	267.74	278.18	10.43	15.3	10.3	-5.0
16	96	15	7557.8	0.003	295.19	300.21	5.02	12.2	10.1	-2.1
17	102	19	7706.2	0.001	269.54	272.84	3.30	12.6	10.2	-2.4
18	108	19	7811.6	0.001	270.69	280.04	9.35	13.2	8.7	-4.5
Average	NA	NA	NA	NA	280.29	287.70	7.41	13.2	9.3	-3.9

Table 4.3 (Columns 6 to 8) provides a comparison of the drayage costs (unit cost per appointment) associated with the optimal solution of the deterministic and stochastic TAS models. As mentioned, the term drayage cost refers to the cost incurred by the drayage companies as a result of following the appointment schedule produced by the stochastic TAS model or the deterministic TAS model. Column 8 shows the drayage cost difference between the stochastic model and deterministic model. As expected, the cost incurred by drayage companies following the appointment schedule produced by the stochastic TAS model is higher than the cost following the schedule produced by the deterministic TAS model (+7.41 unit-cost per appointment on average). The reason is that the stochastic TAS model considered a wide range of turn times, with some being very high and some being very low while the deterministic model considered only the average turn time. As a result, the assigned time-windows deviate more from the preferred time windows for the stochastic TAS model, and hence, higher cost for the drayage companies. Essentially, the stochastic TAS model produces a more "conservative" schedule with longer turn time built-in. The one key benefit of such a conservative schedule is that it has a higher chance of being feasible should the actual turn time on a given day is higher than the average turn time.

The true benefit of a stochastic TAS model is demonstrated by the results shown in Table 4.3 (Columns 9 to 11). The "drayage cost error" is an average of the difference between the expected and actual drayage costs. It provides a measure of how well the produced appointment schedule perform when there is variation in turn time. The lower this value is, the better the appointment schedule. It can be seen that the drayage cost error of the stochastic TAS model is always lower than that of the deterministic TAS model

(-3.9% on average). This result demonstrates that the schedule produced by the stochastic model is more favorable for drayage companies in the long run. In other words, the stochastic model's schedule with its built-in longer turn time will allow more trucks to meet their assigned appointments. This result can be understood with a simple example. Suppose a drayage firm needs three appointments at the terminal (at 8 AM, 10 AM and noon). The stochastic model's schedule is conservative, and thus, assigned appointments at: 8 AM, noon, and 4 PM. With such a schedule, the drayage firm may need to utilize two trucks. On the other hand, the deterministic model's schedule may assign the three appointments at the preferred times. So, the drayage firm will only need one truck. However, due to variation of turn time from day-to-day and within day, there is a high likelihood that the actual turn time is higher than expected, and thus, the truck will not be able to make its second or third appointment in time. This situation will force the drayage firm to use a second truck. In this example, the drayage firm ended up using two trucks with either the stochastic model's schedule or deterministic model's truck. The advantage of the stochastic model's schedule is that it allows the drayage firm to preplan their entire fleet's tours in advance, whereas the deterministic model's schedule forces the drayage firm to make last minute changes. The ability to plan ahead is what allows the drayage firm to reduce the drayage cost (due to overall shorter tour length/time).

Table 4.4 (Columns 1 to 5) shows the results for the stochastic TAS model when the ratio of quotas and requested appointments is varied between 1.5:1 to 3.5:1. The results indicated that when the ratio is increased from 1.5:1 to 2.5:1, the objective function value decreased. This is due to more trucks being allowed to enter the terminal in a given time-window. That is, fewer trucks are forced to select a different time-window. When the ratio is increased from 2.5:1 to 3.5:1, the objective function value did not change. This result indicates that for the problem sizes considered, a ratio of 2.5 is sufficiently large to accommodate all requests in any given time window.

Table 4.4 (Columns 6 to 8) shows the drayage cost (unit cost per appointment) when the ratio of quotas and requested appointments is varied between 1.5:1 to 3.5:1. It can be seen that the drayage cost yield by the schedule of the deterministic TAS model decreases as the ratio increases. A similar trend is observed for the stochastic model, except that the drayage cost was unchanged when the ratio is increased from 2.5:1 to 3.5:1, for the reason noted above. The key takeaway from this result is that the stochastic TAS model is more beneficial to utilize at lower ratios. In practice, it would be more desirable to have lower ratios than higher ratios to avoid creating a congestion at the gate and to balance out the workload between the gate and the container yard.

Table 4.4 Results of the stochastic model and drayage cost for different ratios

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Experiment number	N	*Q:RA	Objective function value (unit cost)	Optimality gap	Drayage cost (unit cost per appointment)		
					Deterministic TAS	Stochastic TAS	Difference
19	1	1.5:1	1715.5	0.0005	313.78	318.42	4.64
20	1	2.5:1	1583.0	0.0005	304.22	309.22	5.00
21	1	3.5:1	1583.0	0.0005	302.94	309.22	6.28

* Quotas:Requested appointments

4.6 CONCLUSION

This paper developed a new TAS model that considers the uncertainty of turn time in truck appointment scheduling. It is the first study to consider this operational characteristic in a mathematical optimization framework. The developed TAS model is formulated as a stochastic program and solved using the Sample Average Approximation algorithm. Numerical experiment results demonstrated the benefit of the stochastic TAS model given its lower drayage cost error by 3.9% compared to the deterministic TAS model. This result implies that the schedules produced by the stochastic TAS model are more robust and are able to accommodate a wider range of turn time scenarios. Another key takeaway from the experiment results is that the stochastic TAS model is more beneficial to utilize when the ratio of quotas to requested appointments is lower. Thus, in practice, when this ratio is more likely to be on the lower end, drayage companies would benefit more if the appointment schedule adopts the stochastic approach described in this paper.

This study has few limitations that should be taken into account: 1) it is assumed that trucks will arrive during their assigned appointment time-windows (i.e., it did not account for unexpected delays due to traffic often arise in practice) , 2) the numerical experiment results are based on hypothetical drayage area, and 3) turn time data and derived distributions used in the experiments came from only one container terminal. While is clear from this study that considering turn time uncertainty in developing truck appointment schedules is beneficial; additional research is needed to identify what level of uncertainty would necessitate the need for a stochastic model. The "uncertainty" could be a combination of turn time variation, trucks arriving late for appointments and trucks missing appointments altogether.

CHAPTER 5

CONCLUSION AND FUTURE RESEARCH

Three research studies were presented in this dissertation to address different aspects of Truck Appointment Systems development. The proposed designs and methodologies will make the truck appointment reservations beneficial to both terminal operators and drayage firms.

Future research studies will focus on 1) identifying what level of uncertainty would necessitate the need for a stochastic model. The "uncertainty" could be a combination of turn time variation, trucks arriving late for appointments and trucks missing appointments altogether., and 2) investigating a design of Truck Appointment System to consider same-day appointment requests (dynamic TAS).

REFERENCES

Asadabadi, Ali, and Elise Miller-Hooks. "Co-opetition in enhancing global port network resiliency: A multi-leader, common-follower game theoretic approach." Transportation Research Part B: Methodological 108, 2018: 281-298.

Asgari, Nasrin, Reza Zanjirani Farahani, and Mark Goh. "Network design approach for hub ports-shipping companies competition and cooperation." Transportation Research Part A: Policy and Practice 48, 2013: 1-18.

Azab, A., A. Karam, and A. Eltawil. A simulation-based optimization approach for external trucks appointment scheduling in container terminals. International Journal of Modelling and Simulation,2019:1-18.

Azab, Ahmed E., Ahmed Karam, and Amr B. Eltawil. "A Dynamic and Collaborative Truck Appointment Management System in Container Terminals." In ICORES, pp. 85-95. 2017.

Bill Mongelluzzo. "LA-LB tech's quest: More trucker dual transactions." JOC magazine. Link: https://www.joc.com/trucking-logistics/la-lb-tech-quest-more-trucker-dual-transactions_20180124.html. Accessed: January 27, 2019

Birge, J. R., and F. Louveaux. Introduction to stochastic programming. Springer Science & Business Media, 2011.

Brodrick, Christie-Joy, Timothy E. Lipman, Mohammad Farshchi, Nicholas P. Lutsey, Harry A. Dwyer, Daniel Sperling, S. William Gouse Iii, D. Bruce Harris, and Foy G. King Jr. "Evaluation of fuel cell auxiliary power units for heavy-duty diesel trucks." Transportation Research Part D: Transport and Environment 7, no. 4 (2002): 303-315.

Brown, G.G., Cormican, K.J., Lawphongpanich, S., Widdis, D.B. Optimizing submarine berthing with a persistence incentive. Naval Research Logistics 44, 1997, 301–318.

Brown, G.G., Lawphongpanich, S., Thurman, K.P. Optimizing ship berthing. Naval Research Logistics 41, 1994, 1–15.

Caserta, Marco, Silvia Schwarze, and Stefan Voß. "Container rehandling at maritime container terminals." In Handbook of terminal planning, pp. 247-269. Springer, New York, NY, 2011.

Castilho, B.D., Daganzo, C.F. Handling strategies for import containers at marine terminals. Transportation Research Part B: Methodological 27 (2), 1993, 151–166.

Chen, G., and Z. Yang. Optimizing time windows for managing export container arrivals at Chinese container terminals. Maritime Economics & Logistics 12, no. 1, 2010, 111-126.

Chen, G., G. Kannan, and M. M. Golias. Reducing truck emissions at container terminals in a low carbon economy: proposal of a queueing-based bi-objective model for optimizing truck arrival pattern. Transportation Research Part E: Logistics and Transportation Review 55, 2013b, 3-22.

Chen, G., G. Kannan, and Y. Zhongzhen. Managing truck arrivals with time windows to alleviate gate congestion at container terminals. International Journal of Production Economics 141, no. 1, 2013a, 179-188.

Chen, G., Govindan, K., Yang, Z.Z., Choi, T.M., Jiang, L. Terminal appointment system design by non-stationary M(t)/Ek/c(t) queueing model and genetic algorithm. Int. J. Prod. Econ.,2013c,146,694–703

Chen, Gang, and Liping Jiang. "Managing customer arrivals with time windows: a case of truck arrivals at a congested container terminal." Annuals of Operations Research 244, no. 2, 2016: 349-365.

Chen, T. Yard operations in the container terminal – a study in the unproductive moves. Maritime Policy Management 26, 1999 (1), 27–38.

Chen, X., Zh. Xuesong, and G. F. List. Using time-varying tolls to optimize truck arrivals at ports. Transportation Research Part E: Logistics and Transportation Review 47, no. 6, 2011, 965-982.

Cheung, R.K., Li, C.L., Lin, W. Interblock crane deployment in container terminals. Transportation Science 36, 2002 (1), 79–93.

Christiansen, M., Fagerholt, K., Nygreen, B., Ronen, D. Maritime transportation. In: Barnhart, C., Laporte, G. (Eds.), Transportation. Handbooks in Operations Research and Management Science. North-Holland, Amsterdam, 2007, pp. 189–284. This volume.

Chung, K.H., Ko, C.S., Shin, J.Y., Hwang, H., Kim, K.H. Development of mathematical models for the container road transportation in Korean trucking industries. Comput. Ind. Eng. 2007, 53 (2), 252–262.

Daganzo, C.F. Crane productivity and ship delay in ports. Transportation Research Record, 1989a, 1251, 1–9.

Daganzo, C.F. The crane scheduling problem. Transportation Research Part B: Methodological 23 (3), 1989b, 159–175.

Eagle Marine Services. http://www.eaglemarineservices.com/ems/. Accessed: 11/20/2017

Fleming, M., Huynh, N. and Xie, Y. ‘An agent-based simulation tool for evaluating pooled queue performance at marine container terminals’, Transportation Research Record: Journal of the Transportation Research Board, 2013, Vol. 2330, pp.103–112.

Fortuny-Amat, José, and Bruce McCarl. "A representation and economic interpretation of a two-level programming problem." Journal of the operational Research Society 32, no. 9, 1981: 783-792.

Gabriel, S. H., A. J. Conejo, J. D. Fuller, B. F. Hobbs, and C. Ruiz. "Complementarity modeling in energy markets. New York: Springer.", 2012.

Gao, Jiyao, and Fengqi You. "A stochastic game theoretic framework for decentralized optimization of multi-stakeholder supply chains under uncertainty." Computers & Chemical Engineering 122, 2019: 31-46.

Giuliano, Genevieve, and Thomas O’Brien. "Reducing port-related truck emissions: The terminal gate appointment system at the Ports of Los Angeles and Long Beach." Transportation Research Part D: Transport and Environment12, no. 7, 2007: 460-473.

Google Maps. Port of Los Angeles and Long Beach. Retrieved from https://www.google.com/maps/place/Port+of+Los+Angeles/@34.0709648,-118.3704361,9.35z/data=!4m5!3m4!1s0x80dd37ae652ef523:0xd88f6472b07283a5!8m2!3d33.7395459!4d-118.2599686. Accessed: 11/22/2017

Guan, C., and R. Liu. Modeling gate congestion of marine container terminals, truck waiting cost, and optimization. Transportation Research Record: Journal of the Transportation Research Board 2100, 2009a, 58-67.

Guan, C., and R. R. Liu. Container terminal gate appointment system optimization. Maritime Economics & Logistics 11, no. 4, 2009b, 378-398.

Guan, Y., Cheung, R.K. The berth allocation problem: Models and solution methods. OR Spectrum 26 (1), 2004, 75–92.

Harrison, R., N., Hutson, J., Prozzi, J., West, J., Gonzalez, and J., McCray. Drayage Activity in Texas. Center for Transportation Research, The University of Texas at Austin, 2007.

Heilig, L., E., Lalla-Ruiz, and S. Voß. Multi-objective inter-terminal truck routing. Transportation Research Part E: Logistics and Transportation Review 106, 2017a, 178-202.

Heilig, L., E., Lalla-Ruiz, S., Voß. port-IO: an integrative mobile cloud platform for real-time inter-terminal truck routing optimization. Flexible Services and Manufacturing Journal 29, 2017b, no. 3-4: 504-534.

Heilig, L., S., Voß. Inter-terminal transportation: an annotated bibliography and research agenda. Flex. Serv. Manufact, 2017, 29 (1), 35–63.

Hill, L. Bruce, Barbara Warren, and Joe Chaisson. An analysis of diesel air pollution and public health in America. Clean Air Task Force, 2005.

Huynh, N. Reducing truck turn times at marine terminals with appointment scheduling. Transportation Research Record: Journal of the Transportation Research Board 2100, 2009, 47-57.

Huynh, N., and C. M. Walton. Robust scheduling of truck arrivals at marine container terminals. Journal of transportation engineering, 2008, 134:347-353.

Huynh, N., and C. Walton, M. Methodologies for Reducing Truck Turn Time at Marine Container Terminals. Research Report SWUTC /05/167830-1, 2005.

Huynh, N., D. Smith, and F. Harder. Truck Appointment Systems: Where We Are and Where to Go From Here. Transportation Research Board 95th Annual Meeting, no. 16-5663, 2016.

Islam, Samsul. "Simulation of truck arrival process at a seaport: evaluating truck-sharing benefits for empty trips reduction." International Journal of Logistics Research and Applications 21, no. 1 (2018): 94-112.

Jiang, Xinjia, Loo Hay Lee, Ek Peng Chew, Yongbin Han, and Kok Choon Tan. "A container yard storage strategy for improving land utilization and operation efficiency in a transshipment hub port." European Journal of Operational Research 221, no. 1, 2012: 64-73.

Jovanovic, Raka. "Optimizing Truck Visits to Container Terminals with Consideration of Multiple Drays of Individual Drivers." Journal of Optimization 2018 (2018).

Kim, K.H. Evaluation of the number of rehandles in container yards. Computers & Industrial Engineering 32 (4), 1997, 701–711.

Kim, K.H., Moon, K.C. Berth scheduling by simulated annealing. Transportation Research Part B: Methodological 37 (6), 2003, 541–560.

Lai, K.K., Lam, K. A study of container yard equipment allocation strategy in Hong Kong. International Journal of Modeling & Simulation 14 (3), 1994, 134–138.

Lai, K.K., Leung, J.W. Analysis of yard crane deployment strategies in a container terminal. In: Proceedings of ICC & IE Conference, Kyungju, Korea, 1996, pp. 1187–1190.

Lai, K.K., Shih, K. A study of container berth allocation. Journal of Advanced Transportation 26 (1), 1992, 45–60.

Lange, Ann-Kathrin, Kristof Ole Kühl, Anne Kathrina Schwientek, and Carlos Jahn. "Influence of drayage patterns on truck appointment systems." In Hamburg International Conference of Logistics (HICL) 2018, pp. 41-59. epubli, 2018.

Li, N., G. Chen, K. Govindan, and Z. Jin. Disruption management for truck appointment system at a container terminal: A green initiative. Transportation Research Part D, 2018: 61:261-273.

Lim, A. The berth planning problem. Operations Research Letters 22 (2–3), 1998, 105–110.

Macharis, Cathy, and Yvonne M. Bontekoning. "Opportunities for OR in intermodal freight transport research: A review." European Journal of operational research 153, no. 2, 2004: 400-416.

Máhr, T., J. Srour, M. de Weerdt, and R. Zuidwijk. Can agents measure up? A comparative study of an agent-based and on-line optimization approach for a drayage problem with uncertainty. Transportation Research Part C: Emerging Technologies 18, no. 1, 2010: 99-119.

Marković, N., Ž. Drobnjak, and P. Schonfeld. Dispatching trucks for drayage operations. Transportation Research Part E: Logistics and Transportation Review 70, 2014: 99-111.

Minh, Chu Cong, and Nathan Huynh. "Optimal design of container terminal gate layout." International Journal of Shipping and Transport Logistics 9, no. 5, 2017: 640-650.

Morais, Philippe, and Elisabeth Lord. Terminal appointment system study. No. TP 14570E. 2006.

Namboothiri, R., and A. L. Erera. Planning local container drayage operations given a port access appointment system. Transportation Research Part E: Logistics and Transportation Review 44, no. 2, 2008, 185-202.

Park, K.T., Kim, K.H. Berth scheduling for container terminals by using a subgradient optimization technique. Journal of the Operational Research Society 53 (9), 2002, 1049–1054.

Park, Y.M. Berth and crane scheduling of container terminals. Department of Industrial Engineering, Pusan National University, Pusan, Korea, 2003.

Park, Y.M., Kim, K.H. A scheduling method for berth and quay cranes. OR Spectrum 25 (1), 2003, 1–23.

Phan, M., and K. H. Kim. Negotiating truck arrival times among trucking companies and a container terminal. Transportation Research Part E: Logistics and Transportation Review 75, 2015, 132-144.

Phan, M., and K., H., Kim, Collaborative truck scheduling and appointments for trucking companies and container terminals. Transportation Research Part B: Methodological 86, 2016, 37-50.

PierPass http://www.pierpass.org/tag/port-of-long-beach/. Accessed: 11/22/2017

Port of Long Beach (POLB) http://polb.com/. Accessed: 11/22/2017

Port of Los Angeles https://www.portoflosangeles.org/maritime/stats.asp. Accessed: 11.22.17

Port of Los Angeles https://www.portoflosangeles.org/maritime/stats.asp. Accessed: 11/22/2017

Ramírez-Nafarrate, Adrián, Rosa G. González-Ramírez, Neale R. Smith, Roberto Guerra-Olivares, and Stefan Voß. "Impact on yard efficiency of a truck appointment system for a port terminal." Annals of Operations Research 258, no. 2, 2017: 195-216.

Santoso, T., S. Ahmed, M. Goetschalckx, and A. Shapiro. A stochastic programming approach for supply chain network design under uncertainty. European Journal of Operational Research 167, no. 1, 2005: 96-115.

Saxe, H., and T. Larsen. "Air pollution from ships in three Danish ports." Atmospheric environment 38, no. 24, 2004: 4057-4067.

Schulte, F., E. Lalla-Ruiz, R. G. González-Ramírez, and S. Voß. Reducing port-related empty truck emissions: a mathematical approach for truck appointments with collaboration. Transportation Research Part E: Logistics and Transportation Review 105, 2017, 195-212.

Schulte, F., R. G., González, and S. Voß. Reducing port-related truck emissions: Coordinated truck appointments to reduce empty truck trips. In International Conference on Computational Logistics, Springer, Cham, 2015, pp. 495-509

Shapiro, A. and T. Homem-de-Mello. On the rate of convergence of optimal solutions of Monte Carlo approximations of stochastic programs, SIAM Journal on Optimization 11, 2000: pp. 70–86.

Shiri, S., and N., Huynh. Assessment of US chassis supply models on drayage productivity and air emissions. Transportation Research Part D: Transport and Environment, 2017.

Shiri, S., and N., Huynh. Optimization of drayage operations with time-window constraints. International Journal of Production Economics 176, 2016, 7-20.

Shiri, S., M. Ng, and N. Huynh. Integrated drayage scheduling problem with stochastic container packing and unpacking times. Journal of the Operational Research Society 70, no. 5, 2019: 793-806.

Taleb-Ibrahimi, M., Castilho, B., Daganzo, C.F. Storage space vs handling work in container terminals. Transportation Research Part B: Methodological 27 (1), 1993, 13–32.

Torkjazi, M., P. S. Mirjafari, and H. Poorzahedy. Reliability-based network flow estimation with day-to-day variation: A model validation on real large-scale urban networks. Journal of Intelligent Transportation Systems, 2017, 1-23.

Torkjazi, Mohammad, Nathan Huynh, and Samaneh Shiri. "Truck appointment systems considering impact to drayage truck tours." Transportation Research Part E: Logistics and Transportation Review 116, 2018: 208-228.

Van der Heyden, W.P.A., Ottjes, J.A. A decision support system for the planning of the workload of a grain terminal. Decision Support Systems 1, 1985, 293–297.

Watanabe, Yutaka. "Evaluation of carbon dioxide emissions from container ports." Journal of International Logistics and Trade 2, no. 1, 2004: 85-93.

Yang, Huiyun, Li Wang, Qi Xu, and Zhihong Jin. "Collaborative optimization of container allocation and yard crane deployment based on truck appointment system." In

Journal of Physics: Conference Series, vol. 1074, no. 1, p. 012185. IOP Publishing, 2018.

Yi, Sanghyuk, Bernd Scholz-Reiter, Taehoon Kim, and Kap Hwan Kim. "Scheduling appointments for container truck arrivals considering their effects on congestion." Flexible Services and Manufacturing Journal, 2019: 1-33.

Zehendner, E., and D. Feillet. Benefits of a truck appointment system on the service quality of inland transport modes at a multimodal container terminal. European Journal of Operational Research 235, no. 2, 2014, 461-469.

Zhang, C., Wan, Y.W., Liu, J., Linn, R.J. Dynamic crane deployment in container storage yards. Transportation Research Part B: Methodological 36 (6), 2002, 537–555.

Zhang, R., W. Y. Yun, and H. Kopfer. Heuristic-based truck scheduling for inland container transportation. OR spectrum 32, no. 3, 2010, 787-808.

Zhang, X., Q. Zeng, and W. Chen. Optimization model for truck appointment in container terminals. Procedia-Social and Behavioral Sciences, 2013,96:1938-1947.

Zhang, Yang, and Ruiyou Zhang. "Appointment of container drayage services: A primary literature review." In 2017 International Conference on Service Systems and Service Management, pp. 1-5. IEEE, 2017.

Zhao, W., and A. V. Goodchild. The impact of truck arrival information on container terminal rehandling. Transportation Research Part E: Logistics and Transportation Review 46, no. 3, 2010, 327-343.

APPENDIX A

OPTIMAL SOLUTIONS OF SIMPLE EXPERIMENTS WITH TWO, THREE AND FOUR DRAYAGE FIRMS

(1)		Experiment number	6	9	11
(2)	*Unrestricted* drayage firms solution	Desired arrival time-windows for every truck for drayage firm 1	[1,10] [4,10,10]	[1,7,7] [2,2,7]	[1,9,9] [3,10,10] [6,6,10]
(3)		Desired arrival time-windows for every truck for drayage firm 2	[10,10] [2,10,10]	[1] [1,7,7] [1,1,10,10]	[2,2,7] [2,2,10,10]
(4)		Desired arrival time-windows for every truck for drayage firm 3	NA	[1] [1,9,9] [1,1,10,10]	[1,9] [5,10,10] [4,4,10,10]
(5)		Desired arrival time-windows for every truck for drayage firm 4	NA	NA	[9] [4,10,10] [7,7,10,10]

(6)	TAS solution	Adjusted arrivals per time-window for drayage firm 1	(1,0,1,0,0,0,0,0,2,1)	(0,3,0,0,0,0,1,2,0,0)	(1,0,1,0,0,2,0,0,3,2)
(7)		Adjusted arrivals per time-window for drayage firm 2	(0,1,0,0,0,0,0,1,1,2)	(2,2,0,0,0,0,0,2,0,2)	(0,4,0,0,0,0,1,0,2,0)
(8)		Adjusted arrivals per time-window for drayage firm 3	NA	(3,0,1,0,0,0,0,0,2,2)	(1,0,0,2,1,0,0,0,3,2)
(9)		Adjusted arrivals per time-window for drayage firm 4	NA	NA	(0,0,0,1,0,0,2,1,0,4)
(10)	*Restricted* drayage firm solution	New arrival time-windows for every truck for drayage firm 1	[1,10] [3,9,9]	[2,8,8] [2,2,7]	[1,9,9] [3,10,10] [6,6,9]
(11)		New arrival time-windows for every truck for drayage firm 2	[8,9] [2,10,10]	[2] [2,8,8] [1,1,10,10]	[9,9] [2,2] [2,2,7]

(12)	New arrival time-windows for every truck for drayage firm 3	NA	[3] [1,9,9] [1,1,10,10]	[9] [4,4] [5,10,10] [1,9,9]
(13)	New arrival time-windows for every truck for drayage firm 4	NA	NA	[8] [4,10,10] [7,7,10,10]